5-5-2005

Intersectoral Size Differences and Migration: Kuznets Revisited

Nejat Anbarci
Department of Economics, Florida International University

Mehmet A. Ulubasoglu
School of Accounting, Economics and Finance, Deakin University, Melbourne, Australia

Intersectoral Size Differences and Migration: Kuznets Revisited*

Nejat Anbarci[†] Mehmet A. Ulubaşoğlu[‡]

May 2005

Abstract

The empirical evidence on the Kuznets hypothesis ranges from positive or negative support to insignificant relationships. Most studies typically try this hypothesis in domains different than the one conceived by Kuznets, which pertains to the industrialization-led urbanization (i.e., significant rural-urban migration) phase of societies. In this paper, we offer a specific channel on Kuznets' hypothesis in his suggested domain. First, we establish theoretically that intersectoral urban-rural size differences result in an intersectoral income inequality, increasing the national inequality. This, in turn, prompts an intersectoral migration, which works as an equilibriating mechanism in the economy, decreasing the inequality in due course. We then successfully test the predictions of the model. The theoretical predictions yield a recursive triangular system, in which we test, i) how the sectoral size differences influence the agricultural income, ii) how a change in agricultural income acts on migration, and iii) what happens to the income distribution as a result of migration. We find a very strong support for the theoretical predictions and the Kuznets hypothesis in its own domain.

JEL Classifications: C51, O14, O15.

*We appreciate useful correspondences and comments on this ongoing project on 'industrialization and migration' with Kaushik Basu, Nancy Chau, Gary Fields, Ira Gang, Cem Karayalcin, M. Ali Khan, Mustafa Caglayan, Naci Mocan and Debraj Ray. We have benefited from useful discussions with Phillip Hone, Dimitrios Thomakos and Maria Willumsen. We also thank the session participants in the 19^{th} Annual Congress of the European Economic Society, held on 20-24 August, 2004 in Madrid, Spain. Any remaining errors are our own.

[†]Corresponding author: Department of Economics, Florida International University, Miami, Florida. e-mail: anbarcin@fiu.edu.

[‡]School of Accounting, Economics and Finance, Deakin University, Melbourne, Australia. e-mail: maulubas@deakin.edu.au.

1 Introduction

Kuznets' seminal contribution made the following observation. Suppose that the urban sector becomes more productive than before and consequently the income gap between rural and urban sectors becomes substantial. Then, inequality in the society would first rise. But as people move from the rural sector to the urban sector as a response to income differential, inequality would then fall:

> "Economic growth perforce brings about a decline in the relative position of one group after another - of farmers, of small scale producers, of landowners." (Kuznets (1973, p.252).
>
> "[O]nce the early turbulent phases of industrialization and urbanization had passed, a variety of forces converged to bolster the economic position of the lower-income groups." (Kuznets (1955, p. 17)
>
> "[A] long swing in the inequality characterizing the secular income structure: widening in the early phases of economic growth when the transition from the pre-industrial to the industrial civilization was most rapid; becoming stabilized for a while; and then narrowing in the latter phases." (Kuznets (1955, p. 18).

In the 1950s, a sufficiently long and reliable time series was available only for the U.S., England and Germany. Kuznets' investigation of these time series indicated that, after having risen earlier, inequality was indeed falling in these countries. Consequently, the Kuznets curve became one of the most heralded stylized facts of the income inequality studies. The common practice in nearly every inequality study has been to test this hypothesis by looking at the sign of a per capita income variable and its square in the inequality regression. The empirical evidence using this approach ranges from positive or negative support to insignificant relationships. Casual model specifications, diversity of the countries used in the analyses and numerous different control variables used seem to have affected the sign of per capita income. Most of these studies essentially checked the behavior of the income distribution with respect to the changes in the 'mean' of the distribution. However, the channels through which this change in the mean can affect the distribution itself have largely been ignored.[1] In sum, these studies offered some results on the "realized" behavior of income inequality within a country over time, virtually independent of Kuznets' own explanation about the pattern. With such an approach, for instance, a very large World Bank data set presented by Deininger and Squire (1996) (for individual countries over time as well as across 108 countries) yielded no support for the hypothesis for more than 95 countries.

It is, however, not surprising that trying the Kuznets hypothesis in domains to which it does not belong does not validate it. As Kuznets (1955) states, "when industrialization and urbanization were proceeding apace and the urban population was being swelled, and fairly rapidly by immigrants... from the country's agricultural areas ... the urban population would run the full gamut from low-income positions of recent entrants

[1] There are, however, some studies which proposed channels to explain the inverted-U shaped pattern of inequality. See, among others, Williamson (1985), Lindert (1986), Aghion and Bolton (1997) and Acemoglu and Robinson (2002).

to the economic peaks of the established top-income groups." The above quote clearly indicates that the Kuznets hypothesis should not be tested for *any arbitrary* period of *any* country. It pertains to the industrialization-led urbanization (i.e., significant rural-urban migration) phase of societies by construction. Additionally, Kuznets does not simply link per capita GDP[2] to inequality in his arguments, as, for him, economic development is a process that entails different phases of industrialization and urbanization, rather than merely being growth in per capita GDP.[3] It is also worth noting that, during the time Kuznets formulated his insights about industrialization and urbanization, developing countries relied heavily on import substitution policies; in a great majority of these countries, import-substitution policies lasted until at least a few decades ago (see Rodrik (1998, 1999a, 2001)).

Given this emphasis on the industrialization-led urbanization and on the specific characteristics of the domain Kuznets specified, one can then wonder whether migration theory can provide any help in explaining the hypothesis. In order to propose an answer to that question in this paper, we first establish an endogenous migration theory, and then use our theoretical results to offer a specific channel and explanation for the inverted-U shape pattern in income inequality. In particular, we establish that as the average manufacturing firm size measured in terms of the number of its workers[4] rises relative to that of an average farm – a very common course in the development stage as well as a prominent feature of the import substitution regimes – intersectoral income inequality increases directly and indirectly (the latter through a consumption channel pointed out by Kuznets). This in turn prompts an intersectoral migration, which works as an equilibriating mechanism in the economy, decreasing later the inequality.

Our theoretical predictions imply fully testable relationships. Using a rich data set, we thoroughly test these predictions and find very strong support for the suggested channel in Kuznets' own domain. In particular, we specify a recursive triangular system in which the channels through which a change in manufacturing firm size relative to that in agriculture affects migration and income equality. This boils down to the estimation of a mechanism in which we test: i) how the intersectoral firm size differences influence - directly and indirectly - the agricultural income vis-a-vis the urban incomes, ii) how a change in agricultural income acts on migration, and iii) what happens to the income distribution as a result of migration. This system is estimated with single-equation and simultaneous equations methods. The estimation results verify the predicted signs for the theoretical variables with high significance levels. These results are relatively robust to many sensitivity checks such as different specifications, regressors, estimation methodologies and measures.

We provide a review of migration literature in Section 2. Section 3 presents the

[2] Kuznets indicates his disapproval of measures such as GDP (Kuznets (1973), pp. 257-8): "[t]he conventional measures of national product and its components do not reflect many costs of adjustment in the economic and social structures to the channeling of major technological innovations; and indeed, also omit some positive returns."

[3] "In view of the importance of industrialization and urbanization in the process of economic growth, their implications for trends in the income distribution should be explored." (Kuznets (1955, p. 12)).

[4] Many studies take the number of workers as the proxy for the size of enterprises; not surprisingly, there are large size differences among countries. Tybout (2000, p. 15) states that "the contrast between the size distribution of plants in developing countries and that found in the OECD is dramatic. ... there is a large spike in the size distribution for the size class 1-4 workers, and it drops off quickly in the 10-49 category among the poorest economies."

model and the predictions on Kuznets' hypothesis; Section 4 describes the empirical methodology. In Section 5 we discuss the data issues; in Section 6 we present our empirical results. In Section 7 we discuss the sensitivity of our results, and Section 8 concludes.

2 Review of Migration Literature

Let us first review the two basic strands of migration literature. The first one is the Lewis framework. The Lewis (1954) model (which was later extended and formalized by Ranis and Fei (1961)) became the dominant paradigm in the second part the 1950s and in the 1960s. The Lewis framework considered internal migration as a natural and desirable process in which surplus labor was gradually shifted from the agriculture to provide needed manpower for urban sector. The process was deemed socially beneficial because human resources were being shifted from the sector where their social marginal product was often assumed to be virtually zero to the sector where this marginal product was significantly positive and rapidly growing. The Lewis model does not explicitly consider utility (or expected earnings) maximization of agents. Thus, with its ad hoc structure, it fails to throw much light on the migration decision and the functioning of labor markets etc. In addition, as Ray (1998, p. 367) states, the Lewis model implicitly assumes that family farms are taxed as labor is withdrawn from agriculture, keeping agricultural per capita income constant and allowing the supply curve of labor to industry to remain perfectly elastic.

The Harris-Todaro (H-T) framework (i.e., Todaro (1969) and Harris-Todaro (1970) models as well as subsequent work that extended these models), elaborated many features of migration that went unnoticed within the general and ad hoc structure of the Lewis framework. The H-T framework (which is a clear improvement over the Lewis framework in many ways) became the mainstream paradigm in the internal migration literature since the early 1970s. It postulated that migration proceeds in response to urban-rural differences in expected earnings rather than actual earnings. In that framework, given a politically-determined minimum urban wage that exceeds the agricultural wage, the probability of getting a formal sector job is determined by the ratio of available formal sector jobs to job seekers. Consequently, in equilibrium, the expected formal sector wage (i.e., wage times the probability of finding such a job) is equal to the rural wage.

It is not difficult to see that the core assumption of the H-T framework, 'the presence of an exogenous urban-rural income gap', makes that framework problematic both at the empirical[5] and theoretical levels. Consequently, there has been some notable attempts to provide an endogenous explanation for this urban-rural income gap. The first and most notable one is by Stiglitz (1974). Its basic idea was to introduce labor turnover, which is costly to urban employers. The pace of work in the urban sector may slow down due to turnovers. Suppose that a firm paying a higher wage can face a lower labor turnover rate; one can then explain why urban employers may not lower the wages they pay despite the presence of unemployment.[6]

[5]Several recent studies (for instance, Amsden and van der Hoeven (1996), Hoddinott (1996), and Levy and Newman (1989)) verify that, even though such high sticky urban wages were common in some African countries at the time the H-T framework was formulated, they proved to be temporary.

[6]In another attempt to endogenize the urban-rural gap, Calvo (1978) used a trade union whose objective is to maximize the difference between its members' urban income and what they would get in the rural sector

The papers that endogenized the urban-rural income gap of the H-T framework, however, maintained other exogenous features of the H-T framework. For instance, prices of the agricultural and manufacturing goods are taken as unity in all of the above papers - this being a consequence of their small open economy assumption. This assumption, on the other hand, is not realistic considering that many countries have been experiencing significant internal migration levels since the 1950s, and many of these countries were not exactly open economies especially during that era, relying heavily on import substitution policies as mentioned in the Introduction.[7] In addition, in this literature, agents' labor supply and leisure decisions are typically by-passed; the notable exceptions are Bhatia (1979) and Gang and Gangopadhyay (1987).

Although industrialization-led urbanization (i.e., rural-urban migration) is the engine of the Kuznets hypothesis, it is not possible to find such a micro-foundation for the Kuznets hypothesis in either the Lewis framework or the H-T framework. Another work in migration, Sjaastad (1962), hinted the right starting point by emphasizing 'the influence of migration as an equilibrating mechanism in a changing economy'. Instead of only considering the question as to 'how effective is the urban-rural income gap in explaining migration', Sjaastad rightfully emphasized another very important question, 'how effective is migration in equalizing urban-rural earnings'. Thus, any migration model that can serve as the micro-foundation of the Kuznets hypothesis should explicitly incorporate features that answer the latter question. To this extent, in the next section we will employ a very simple and parsimonious model to illustrate the channels through which Kuznets hypothesis operates.

3 The Model and the Predictions

We use a simple two-sector model in which variables such as labor demand, labor supply and leisure of agents as well as the resulting sectoral prices, manufacturing wage, rural and urban profit are determined endogenously. Thus, the model examines the effects not only of urban-rural income gap on migration but also of migration on the urban-rural income gap. The spirit of this work is the human-capital models (such as Sjaastad's (1962)), which view migration as a personal investment that will be made only if its returns are justified (in human capital models, the benefits of migration occur over a period of time - which also helps explain why migration rates decline in the potential migrant's age).

In our simple framework, all households possess a simple Cobb-Douglas utility function the arguments of which are leisure, an agricultural good and a manufacturing good. There is only one variable input in the economy, namely labor (as in the other papers in the migration literature). Rural family farms, using their own family labor, produce the

(that way a fixed urban-rural income gap is always maintained). Quibria (1988) extended Calvo's model by considering (i) an informal as well as a formal urban sector, and (ii) risk-averse agents; the trade union's objective is to maximize the difference between its members' urban formal sector income and what they would get in the urban informal sector.

[7] Instead of these normalized prices that were in the above papers, the original Harris and Todaro (1970) paper had used an ad hoc price equation where the price of the agricultural good - in terms of the manufacturing good - is a function of the relative outputs of the two goods respectively. Gang and Gangopadhyay (1985) went one step deeper and used a demand equation instead.

agricultural good.[8] In the manufacturing sector, there are entrepreneurs and workers. For simplicity, it is assumed that each firm is owned by one entrepreneur. In addition, each firm's manufacturing technology allows using k times more labor than a typical farm uses.[9] It turns out that this difference between the two sectors' production structures[10] suffices to generate the urban-rural income gap in our model. Farmers respond to this income gap by migrating to the urban manufacturing sector where they become workers.

We will first establish the equilibrium level of leisure and work hours of the agents, wages, and the domestic terms of trade. Using them, we will establish the agents' income levels and consequently their utility levels. It will turn out that the income gap between workers and farmers will increase in the relative firm size of industry to agriculture, k, and decrease in the fraction of agents' income spent on the agricultural good, α. Migration will decrease inequality in due course.

3.1 Agriculture

The agricultural sector consists of small producers, who use their own labor,[11] and experience diminishing marginal product of labor. The amount of land and capital of each farm will be normalized to one unit. A farm's production function is $(l_a)^{\frac{1}{2}}$, where l_a denotes the agricultural labor used in each farm with $l_a \in (0,1)$.[12] Let p_a denote the agricultural good's price.

Thus, a farm's (or in short a farmer's) income is $Y_a = p_a(l_a)^{\frac{1}{2}}$ (observe that $Y_a = p_a(l_a)^{\frac{1}{2}} = \pi_a + w_a l_a$ since $\pi_a = p_a(l_a)^{\frac{1}{2}} - w_a l_a$ where π_a is the agricultural profit and w_a is the agricultural wage rate).[13]

A farmer's utility is given by $u_a = L_a a_a^{\alpha^f} m_a^{1-\alpha^f}$ where L stands for leisure and $L_a = 1 - l_a$ holds, with $\alpha^f \in (0,1)$.[14]

3.2 Manufacturing

At a given point in time, the fraction of labor force who work in the rural sector and produce agricultural output will be denoted by $A \in (0,1)$. Thus, the fraction of labor

[8] Otsuka, Chuma and Hayami (1992, p. 1971, Table 1) report that about 80 percent of farms worldwide are owner-cultivated, and thus are small-sized in that only family members - and at peak seasons a few others - work the land. See Otsuka *et al.* (1992) and Ray (1998, p. 417-419) as to why even 80 percent is an understatement.

[9] A detailed description of k will be given below.

[10] Adam Smith (1937, p. 5-6) stated: "The nature of agriculture, indeed, does not admit of so many subdivisions of labour, nor of so complete a separation of one business from another, as manufactures. It is impossible to separate up entirely, the business of the grazier from that of the corn-farmer. ... The spinner is almost always a distinct person from the weaver; but the ploughman, the horrower, the sower of the seed, and the reaper of the corn, are often the same. ..."

[11] This feature is not commonly used as Otsuka *et al.* (1992, p. 1966) state "the existing models do not explicitly consider ... that the majority of farming households are owner cultivators."

[12] Our production functions in both sectors are slightly more specific forms than the ones used in Calvo (1978) and Quibria (1988). Consequently, these specific functional forms allow us to obtain more explicit expressions in return.

[13] The farmer is a representation of a household farm. "In ... farming, [the] management skill is combined with the farm household's own labor power" (Timmer (1988)).

[14] It is well-known that in Cobb-Douglas utility functions, the portions of income spent on different goods are proportional to the exponents of those goods. Thus, α^i portion of each agent i's income is spent on the agricultural good and $(1-\alpha^i)$ portion is spent on the manufacturing good.

force who work in urban sector and produce manufacturing output will be $1 - A$.

In the manufacturing sector, let $1/(k+1)$ fraction of individuals be entrepreneurs - who own the firms - and thus, $k/(k+1)$ fraction of individuals be workers, where k is an integer greater than 1. (Clearly, without this assumption both sectors would have the same small-producer structure, and thus the same income level by every one.) Thus, there are k workers per firm, and thus per entrepreneur. In other words, k represents the relative firm size of industry to agriculture (since in agriculture implictly k is 1). It is important to note that in this framework, k also stands for the extent of technology in terms of division of labor a la Adam Smith's "pin factory" example.[15] Hence the population fraction of farmers is A, while the population fraction of workers is $k(1-A)/(k+1)$ and that of entrepreneurs is $(1-A)/(k+1)$.

A farmer who migrates to the urban sector becomes a worker (first) [16,17]

Meanwhile, in order to maintain the consistency of the model about k (and, at the same time, to bypass the complications of each period workers competing to become entrepreneurs), we will assume that (i) given positive time preferences and long time horizons, a worker becomes an entrepreneur with a very small probability towards the end of his/her career, and that (ii) k nevertheless keeps increasing as migration takes place with the net effect that the number of entrepreneurs increasing too. (Observe that part (i) of the above assumption will help slow down the increase in k to a small extent). The setup we have in mind is follows: at the beginning of Period 0, k increases due to an exogenous outside factor. Then, if $(1-A)$ increases at the end of each Period t, the next Period t+1 will start with a higher than that of Period t (e.g., as a result of migration k will increase further). Thus, if migration takes place at the end of each self-contained period, the next period will start with a higher k than before. If no migration takes place, k will not increase.

We will normalize the price of the manufacturing good by assuming $p_m = 1$. The profit of the firm, given its Constant-Returns-to-Scale production function $k(l_m)^{1/2}$, is $\pi_m = k(l_m)^{1/2} - w_m k l_m$ where π_m denotes the firm's profit, w_m is the manufacturing sector's wage, and l_m denotes the labor input by an individual that works in a typical firm in that sector. The entrepreneur draws his profit on one (normalized) unit of capital.[18]

[15] "To take an example, ... a workman ... could scarce, perhaps, with his utmost industry, make one pin in a day, and certainly could not make twenty. [With division of labor,] [o]ne man draws out the wire, another straights it, a third cuts it, a fourth points it, a fifth grinds it at the top for receiving, the head; [i]n this manner, divided into about eighteen distinct operations, which, in some manufactories, are all performed by distinct hands, ... (they can make) forty-eight thousand pins in a day. " (A. Smith 1776)

[16] After citing a major survey, Cole and Sanders (1985, p. 482) state that " the mean level of education of urban immigrants had regressed toward the rural mean. ... Migration is becoming less and less selective. ... [Immigrants] move to the city with the expectation of finding long-term employment" at a low paying job. Thus, this precludes the possibility that some migrants' becoming entrepreneurs instantly upon their arrival at the urban sector.

[17] An important question is 'what happens to the land that is left behind the migrant'. In general two things happen. First, due to high population growth in rural areas, per capita land decreases. Consequently, especially the migrants' relatives who are left behind take over that land; this feature is not part of our model. The second and maybe less obvious explanation would be that in time some agricultural land around many cities keeps joining the urban land of those cities; this feature is fully reconcilable with our model.

[18] Apart from the obvious simplifying advantages that this assumption buys us, (as a footnote in Section 3.6 alludes to) the interruptions experienced by many developing countries in the imports of the essential capital goods provide another justification for this assumption. There are, however, other justifications too for limiting the amount of capital to a fixed level. As Tybout (2000, p. 17) states, in LDCs, "[p]lentiful unskilled labor and

An entrepreneur's utility is given by $u_e = L_e a_e^{\alpha^m} m_e^{1-\alpha^m}$ where $\alpha^m \in (0,1)$.

In the above utility specification for entrepreneurs, we use α^m instead of α^f that we have used in the utility function of farmers. Kuznets (1966, p. 102) links technological changes - which are typified by increases in k in our framework a la the pin factory example of Adam Smith - to increases in $(1-\alpha^m)$ for the city dwellers.[19] A convenient and straightforward way of representing that link is to assume $\alpha^m = 1/k$ (and thus,$(1-\alpha^m) = (k-1)/k$). Hence, in the remainder of the paper, we will assume for farmers and $\alpha^m = 1/k$ for entrepreneurs and workers (where the latter two live and work in the urban sector).

Hence, an entrepreneur's income is $Y_e = \pi_m + w_m l_e$.

Now consider a worker's utility function, which is given by $u_w = L_w a^{\alpha^m} m^{1-\alpha^m}$ where $L_w = 1 - l_w$. A worker's income is $Y_w = w_m l_w$.

3.3 Equilibrium Levels of Incomes

Our first result provides the equilibrium levels of income for farmers, entrepreneurs and workers, and how the levels of α^f, k and A affect these equilibrium incomes.

Proposition 1: (1) A farmer's income is $Y_a^* = [\frac{1-A}{A}\frac{1}{(1-\alpha^f)(k+1)}](\frac{1}{2})^{1/2}$.
(2) An entrepreneur's income is $Y_e^* = \pi_m^* = k\frac{1}{2}(\frac{1}{2})^{1/2}$.
(3) A worker's income is $Y_w^* = l_m^* w_m^* = \frac{1}{2}(\frac{1}{2})^{1/2}$.

Due to our normalization $p_m = 1$, a worker's income Y_w^* comes out fixed at a normalized level $\frac{1}{2}(\frac{1}{2})^{1/2}$. As a result, the incomes of an entrepreneur and a farmer are expressed relative to that of a worker. Observe that an entrepreneur's income is k times that of a worker.

For agrarian societies, A is very close to 1 (and thus $(1-A)$ is very close to 0). Since $k > 1$, even for very high levels of α^f, it is straightforward to see that $Y_w^* > Y_a^*$ should hold for such agrarian societies (this is because $(1-A)/A$ and consequently Y_a^* will be very small). But for very low levels of A (i.e., in a highly urbanized society), $Y_w^* > Y_a^*$ need not hold (this is because $(1-A)/A$ and consequently Y_a^* will be very high). Thus, the equilibrium income levels found in Proposition 1 lead us to the intersectoral income gap.

the lack of long-term finance create incentives to economize on fixed capital" and "[s]ince most machinery and equipment must be imported, the trade regime and the lack of local technical support" may further add to these incentives. In addition, "[i]f there is substantial uncertainty about future demand conditions ..., it often makes sense to choose production techniques that do not lock in: that is, to rely more heavily on labor" (Tybout (2000, p. 18)).

[19] Kuznets (1966, p. 102) establishes a link between k and α^m: " [t]he technological changes that have accounted for the rise of modern industry and for concomitant industrialization and urbanization have meant that consumers, who as producers had to live in the cities, have required goods and services that were not essential in the countryside. ... television sets can be seen as a substitute for village feasts in the way of recreation, and railroads and automobiles as a substitute for horses in the way of transportation. Even if the structure of wants in their distribution among broad categories such as nourishment, clothing, shelter, recreation, transportation, and the like, had not changed - in the sense that the proportions of total product allocated to each remained constant - with the differential effect of technological innovations the same wants could have been satisfied in one period by products from the agriculture sector and in another by products from the industry sector."

Corollary 1: Y_w^*/Y_a^* and Y_e^*/Y_a^* (and thus $Y_w^* - Y_a^*$ and $Y_e^* - Y_a^*$) increase in k and A but decreases in α^f.

That is, intersectoral income ratio (and gap) increases as there is more labor per entrepreneur, as there is less urban labor force, and as there is less demand for agricultural goods, because these factors lower a farmer's income while they do not affect the income of a worker - and an increase in k increases an entrepreneur's income. In particular the income gap between workers and farmers will be crucial in determination of migration, which is the response of farmers to this income gap. A brief review of the proof of Proposition 1 will reveal that Y_w^*/Y_a^* and Y_e^*/Y_a^* (and $Y_w^* - Y_a^*$ and $Y_e^* - Y_a^*$) depend on k both directly - and through $\alpha^m = 1/k$ - indirectly.

3.4 Equilibrium Level of Relative Labor Force

Migration decision of a farmer - who becomes a worker (first) upon migrating to the urban sector -, however, does not depend on whether or not Y_w^* exceeds Y_a^*. It depends on whether or not a farmer's utility u_a^* exceeds a worker's utility u_w^*. In other words, migration in equilibrium will cease when the utility from migrating is less than the utility of not migrating for the marginal potential migrant. The first part of the theorem below states that migration stops once A tapers off to a certain level. To shed light on the effects of α and k on migration, the second part of the theorem considers a special yet reasonable and tractable case - i.e., the case where α^f and α^m tend to each other. Given that special case, we are able to identify the circumstances under which migration will continue and how α and k affect the equilibrium fraction of urban population. In sum, a high k and low α will cause a higher level of urban population:

Theorem 1 (1) An equilibrium level of $\frac{1-A}{A} \in (0,1)$ exists.

(2) Suppose α^f and α^m tend to each other such that in the limit $\alpha^f = \alpha^m = \alpha$. Then migration will continue as long as $\frac{1-A}{A} < \frac{(1-\alpha)(k+1)}{2}$ and will stop when $\frac{1-A}{A} = \frac{(1-\alpha)(k+1)}{2}$. In addition, the equilibrium level of $\frac{1-A}{A}$ will decrease in α and increase in k.

Part 1 of Theorem 1 implies that at some A^*, u_a^* will reach the level of u_w^* - despite the presence of differing α^f and α^m - and migration will stop. Differences in any two agents' utilities may be due to three factors: incomes, leisure levels and the αs. For instance, by the proof of Proposition 1 in the Appendix, an entrepreneur's leisure turns out to be higher than those of workers and farmers. However, the leisure levels of farmers and workers turn out to be the same. When the incomes of workers and farmers become equalized, the only difference between the utilities of those agents can be the αs. But, as the second part of Theorem 2 assumes, as α^f and α^m tend to each other, the utilities of farmers and workers will be equalized and migration will stop. The condition that makes their incomes equal (in the presence of convergence between the αs) is $\frac{1-A}{A} = \frac{(1-\alpha)(k+1)}{2}$.

Note that Theorem 1 provides the *equilibrium conditions* among A (hence $\frac{1-A}{A}$), α and k, which are the snapshots of the economy at some particular points in time. The income equilibrium condition in Theorem 1, $\frac{1-A}{A} = \frac{(1-\alpha)(k+1)}{2}$, provides the circumstances under which migration stops. In general it is conceivable that α and k will keep changing constantly and A will keep adjusting to the changes in α and k. What levels of α and k would lead to a high or low equilibrium levels of A (or of $\frac{1-A}{A}$)? Observe that, if in equilibrium $\frac{1-A}{A}$ is realized at a rather low level, it means that migration has stopped when the urban population was still at a rather low level. For that to happen, k must be rather low and α must be rather high. On the other hand, if in equilibrium $\frac{1-A}{A}$ is realized at a rather high level, it means that migration has stopped when the urban population has already reached a rather high level. For that to happen, k must be rather high and α must be rather low.

3.5 Equilibrium Level of Inequality

An important question is 'what affects *inequality in the society* in what direction during industrialization'. Corollary 1 above indicates that the *intersectoral income inequality* will widen as k increases and α decreases, but will shrink in response to a decrease in A (i.e. migration). But note that Corollary 1 does not touch upon the fact that an increase in k also widens the income ratio (and gap) between entrepreneurs and workers. To measure the level of "overall" (i.e., not just intersectoral) inequality in the society, we will use the Gini coefficient, G. A higher Gini coefficient corresponds to a higher income inequality. Our next theorem verifies that, as is the case in the intersectoral inequality, a high intersectoral firm size ratio, k, and low agricultural consumption fraction, α, will increase inequality in the society, and a higher urban population fraction (i.e., a higher accumulated migration), $(1-A)$, will decrease it.

Theorem 2 G decreases in $(1-A)$ and α, but increases in k.[20]

To see the intuition of this result, consider Corollary 1 (as well as the specific results of Proposition 1). As k increases, the incomes of farmers will fall below those of workers and (especially) entrepreneurs. Likewise, as α decreases, the incomes of farmers will fall further below those of workers and entrepreneurs. But, as $(1-A)$ increases (i.e., as A decreases) - which is only possible through migration -, the incomes of farmers this time will get closer to those of entrepreneurs and workers.

[20] One might consider relating G to α and k only since Theorem 1 relates A to α and k. But recall that Theorem 1 states only the equilibrium conditions among A, α and k. That is, as noted before, in general α and k may keep changing and A will have to keep adjusting to the changes in α and k. Hence, Theorem 2 does not only consider inequality at equilibrium levels of A, α and k only; rather, it considers inequality given all possible levels of A, α and k.

3.6 Model's Relevance to the Kuznets Hypothesis

By utilizing Kuznets' (1966) observation on the negative relationship between α and k as well as our Theorems 1 and 2, a comparative-static setup (which entails k, α and A) can illuminate the relevance of our analysis for the inverted-U shape hypothesis on inequality. We provide the steps for this setup below.

1. An increase in k **increases intersectoral and national inequality directly and indirectly**.[21] It was observed in the past that k increased over the course of industrialization, with especially the contributions of interruptions in the capital goods imports.[22] The direct effect of k on inequality is that inequality rises by Corollary 1, as the incomes of farmers fall below those of workers and (especially) of entrepreneurs. Inequality rises further with the indirect consumption effect: the increase in k decreases α^m (in reality, as mentioned above, it may even decreases α^f) – i.e., an absolute fall in farmers' income. In a society, although α^f can be significantly different than α^m, the two cannot stay too far apart form each other for too long. In addition, due to the data limitations (since we do not have an α^f and α^m split in the available data) - and as our assumption in Part 2 of Theorem 1 -, we will use one common α instead of separate α^f and α^m.

2. **Migration increases in response**. Part 2 of Theorem 1 predicts that the (equilibrium) level of urban-rural labor force ratio (where it reaches after migration) depends positively on the level of k and negatively on the level of α. Thus, this implies that the increase in k fosters migration, given the level of α, and the equilibrium urban-rural labor force ratio decreases in α, given the level of k.

3. **Inequality decreases in response to migration**. Theorem 2 predicts that inequality depends negatively on $(1 - A)$ and α, and positively on k. This suggests that an increase in $(1 - A)$ (i.e., accumulated migration) works as an equilibriating mechanism in the economy, and decreases the inequality, given the levels of α and k.

4 Econometric Analysis

The predictions of the theoretical model above implies fully testable relationships. Clearly, similar predictions can be obtained by using different assumptions and approach as well. In our empirical analysis, therefore, we will adhere to a literal test of the theoretical

[21] Although for the sake of the model's parsimony, we have taken the levels of α^f and k fixed, they are hardly fixed in real life.

[22] In many developing countries, capital goods exports (which have been essential for their capital stock enhancements) have been interrupted frequently for extended periods: "During the depression of the late 1930s ... international trade was modest and declining. Interruption of shipping and non-military production during World War II ... imports fell sharply, especially those classified as capital goods." (Bruton (1988, pp. 1611-12).) In addition, in the seventies and eighties, many of these countries which continued their Import Substitution policies were not able to import essential capital goods due to their dismal balance of payments problems. "Yet ... domestic demand was strong. There was a powerful incentive to find ways to increase productivity and output. In these years producers searched for ways to increase output from the resources available within the country. ... Their only asset is their labor, their people, They therefore had to search ways to make their labor more productive." (Bruton (1988, pp. 1612-13).)

predictions both to check the empirical validity of our theoretical assumptions as well as to see if the predictions receive support from data. In particular, we test the Kuznets' observation on the negative relationship between α and k (the indirect effect) and our Theorems 1 and 2 (the direct effect). In doing so, we specify a mechanism and channel through which the inverted-U shape pattern in income inequality emerges.

4.1 Predicted Specification

The econometric model specification predicted by the theory is given below:

$$\alpha_{it} = \delta_0 + \delta_1 k_{it} + \sum_{p=2}^{P} \delta_p X_{it} + u_{it} \tag{1}$$

$$\frac{(1 - A_{it})}{A_{it}} = \gamma_0 + \gamma_1 \alpha_{it} + \gamma_2 k_{it} + \sum_{r=3}^{R} \gamma_r Y_{it} + v_{it} \tag{2}$$

$$INEQ_{it} = \phi_0 + \phi_1 \frac{(1 - A_{it})}{A_{it}} + \phi_2 \alpha_{it} + \phi_3 k_{it} + \sum_{s=4}^{S} \phi_s Z_{it} + \omega_{it} \tag{3}$$

where p, r, s are parameter indices, α is the fraction of income spent on agricultural goods in a country (percentage share of agricultural value added plus agricultural exports minus agricultural imports in GDP), k is the number of workers per entrepreneur in each country (the ratio of number of employees to the number of establishments in manufacturing industry), A is the share of agricultural labor force in the total labor force, thus $\frac{(1-A)}{A}$ is the ratio of non-agricultural to agricultural labor force, and $INEQ$ is the income inequality indicator, i.e., the Gini index. The vectors $\mathbf{X}$, $\mathbf{Y}$ and $\mathbf{Z}$ include the control variables for the respective equations. We describe our benchmark control variables below, and their alternatives are discussed in Section 7.

In Equation (1), we test the inverse relationship between the fraction of income spent on agricultural goods, α, and the intersectoral firm size ratio, k, as offered by Kuznets. Thus, we expect a negative sign for δ_1.

In Equation (2), we test the effects of α and k on the equilibrium urban to rural labor force ratio. An increase in k would make farmers worse off. A decrease in α would intensify this, as the farmers live on agricultural goods. In theory, farmers respond to this decrease in α by migrating to urban manufacturing sector. Thus we expect a negative sign for γ_1 and a positive sign for γ_2.

In Equation (3), the effects of $\frac{(1-A)}{A}$, k and α on inequality are tested. Based on the theoretical premise that rural people will be employed in the manufacturing sector after migration and start earning the same income as initial workers, we expect a higher urban labor force fraction to decrease inequality. This predicts a negative sign for ϕ_1. This process would take place given the levels of k and α, and inequality will be increasing in k and decreasing in α.

A very significant point is that the ascension of the countries over the inverted-U shape of inequality is tested through the impact of α and k on $\frac{(1-A)}{A}$ and $INEQ$, and the descension is tested through the impact of α and $\frac{(1-A)}{A}$ on $INEQ$.

One can suggest another equation into the system that tests the impact of k on the intersectoral income ratio (and/or gap), and capture the ascension along the inverted-U shape in that way. However, this is the prediction of Corollary 1, which ultimately leads to Theorem 1, from where we derive our estimable equations. As per Theorem 1, the sectoral income inequality-increasing effect of k is tested through Equation (2), and we interpret a positive sign for γ_2 as the ascension over inverted-U shape.[23]

4.2 Data and Sample Issues

As a reflection of the idea that the Kuznets hypothesis cannot be tested for any country in any arbitrary period, we test the predictions of the theory across different samples. The model is written in the specific context of a developing country with dual economic structure to echo Kuznets' reasoning on the inverted-U shape. This sort of economic structure in general accords with the *initial* phases of the import-substition regimes, where, as in our model, agricultural and manufacturing prices are endogenously determined in domestic markets. Indeed the theoretical variable k is observed to be rising in this period too. Thus, an empirical test in conformity with the model's predictions should be undertaken with data from the closed-economy periods of the countries. We are able to match the *closed economy*-$\alpha - k - \frac{(1-A)}{A} -$ *Gini* combination of the data for 32 developing countries, providing us with 78-80 usable observations in the estimations.[24] Because this sample brings together a relatively homogenous lot of countries, we call this our "base" sample, and it is thought to be the most suitable sample to find support for the model and Kuznets' reasoning. Note that this sample excludes Sub-saharan African countries, whose economic, social and political progresses have been different than other "normal" developing countries. Indeed 13 Sub-saharan African countries provide the necessary data with 19-21 observations,[25] enabling us to build another sample that combines the base and Sub-saharan African countries. It is expected that the data in this sample will provide less support for the theoretical predictions. Moreover, 12 developed countries provide 15 observations on the necessary variables,[26] giving us the opportunity to build

[23] While we report the results for this approach in this paper, we also exercise with the former and find supporting results for our predictions.

[24] The countries and the periods are: Bangladesh 65-69, 70-74, 75-79, 80-84, 85-89; Bolivia 80-84; Brazil 70-74, 80-84, 85-89; Chile 65-69, 70-74; Colombia 65-69, 70-74, 75-79, 80-84; Costa Rica 80-84; Dominican Republic 75-79, 80-84, 85-89; Ecuador 65-69, 85-89; Egypt 65-69, 75-79, 80-84, 85-89; El Salvador 65-69, 75-79; Fiji 75-79; Guatemala 80-84; Guyana 80-84; Honduras 65-69, 80-84, 85-89; India 80-84, 85-89; Indonesia 60-64, 65-69; Iran 65-69, 70-74, 80-84; Jamaica 75-79, 80-84, 85-89; Mexico 80-84; Morocco 65-69, 80-84; Nepal 75-79, 80-84; Pakistan 65-69, 70-74, 80-84, 85-89; Panama 65-69, 70-74, 75-79, 80-84; Peru 70-74, 80-84, 85-89; Philippines 60-64, 65-69, 70-74, 80-84; South Korea 60-64; Sri Lanka 65-69, 70-74; Sudan 65-69; Trinidad and Tobago 70-74, 75-79, 80-84; Tunisia 65-69, 75-79, 80-84; Turkey 70-74, 75-79, 80-84; Venezuela 70-74, 75-79, 80-84, 85-89.

[25] The countries and the periods are: Burundi 90-94; C. Afr. Rep. 90-94; Cote d'Ivoire 85-89; Ghana 80-84; Kenya 80-84, 85-89; Lesotho 85-89; Nigeria 60-64, 85-89; Senegal 90-94; South Africa 65-69, 80-84, 85-89; Tanzania 65-69, 75-79, 90-94; Uganda 85-89; Zambia 60-64, 75-79, 80-84; Zimbabwe 90-94.

[26] The countries and the periods are: Australia 60-64; Bahamas 75-79, 80-84; Canada 60-64; Finland 60-64; Germany 60-64; Greece 60-64; Ireland 60-64; Japan 60-64; Netherlands 60-64; New Zealand 70-74, 75-79, 80-84; Spain 60-64; Sweden 60-64.

additional two samples: base + developed, and base + developed + Sub-saharan African. These two samples, too, are expected to provide less support for the theory than only the base countries sample would. Therefore we have four samples altogether to undertake the analysis. The specifics of the results, however, may be different in each sample, with a possibility of offering very interesting implications. For instance, we expect different curvatures of the inverted-U shape, if any, to arise in different bunch of countries.

We determine the opening years for each country as designated by Sachs and Warner (1995) and Wacziarg and Welch (2003).[27] The agricultural tastes variable, α, is approximated by the share of agricultural value added in GDP, adjusted for agricultural imports and exports. The share of labor force in agriculture and non-agricultural sectors is a relatively straightforward to handle. We interpret a decrease in the share of agricultural labor force as migration to manufacturing.

Of the key variables in our setup, k is measured by the ratio of number of employees to the number of establishments in manufacturing (recall also from the theoretical model that each establishment is assumed to be owned by one entrepreneur). We have data for k as far back as 1950s and 1960s, and thus its evolution within countries is captured to a great extent. This variable may be considered as very literal to use in the empirical analysis (given that k is defined in the model as the firm size in manufacturing relative to that in agriculture, whose size is assumed to be 1), however the data on the average number of employees provide an interesting rising trend over time for many countries, lending credence to the channel we are offering. Figures for most of our countries are included in the Appendix.[28] As noted above, this increase can be due to the development of labor-intensive primary manufacturing sectors in the initial phases of import-substitution periods, or to the substitution of capital with labor should the countries experience restrictions on capital goods exports, which was observed happenning in the latter periods of the import-substituting regimes. The relevant data on k are available in United Nations UNIDO and General Industrial Statistics databases. In the robustness section, we use the ratio of manufacturing value added to agricultural value added to check the relevance of k in representing the relative size of manufacturing and agricultural sectors across countries as implied by Tybout (2000).

We obtain the inequality data from the United Nations WIDER database, which is an augmented version of the Deininger and Squire (1996) (henceforth D-S) data set with some developed and transition countries data added. Very interestingly, most inequality studies have found conflicting results on the Kuznets hypothesis by using the D-S data.[29]

[27] Sachs and Warner (1995) specify five criteria for a country to be considered as open: i) average tariffs rates being 40% or less, ii) nontariff barriers covering 40% or less of trade, iii) black market premium on the exchange rate being less than 20%, iv) no state monopoly on major exports, and v) not being a socialist system.Wacziarg and Welch (2003) revise the Sachs-Warner classifications for some countries, such as India. Rodriguez and Rodrik (2000) closely and excellently examined the dichotomous openness variable of Sachs and Warner by partitioning it into its original components. Although they show that as an "openness index" this variable is not the most appropriate measure in trade policy-growth context, they mention that most of the variation in this variable is captured by black market premium, which is an indication of poor domestic policies of closed trade regimes. Hence, they conclude, "the Sachs and Warner indicator serves as a proxy for a wide range of policy and institutional differences," the qualifications that we look for in our analysis.

[28] Not every country provides a continuous series that can be presented graphically, although point observations of some countries that are not in the graphics have been used in the estimations.

[29] Depending on the focus of the papers, it is reported that even a slight change in the state variables, control variables, time span or functional form etc. result in different implications about the Kuznets hypothesis.

These results naturally casted doubts on the quality of the data set. Since we have a model and a domain most relevant to Kuznet's own explanation for the hypothesis, the use of D-S data set in this framework will be full of implications.

In addition, we use certain dummy variables to control for differences in the construction of the Gini data. D-S report Gini data from various sources, and classify them acceptable vs. non-acceptable based on certain criteria. Even within the class of acceptable data, there are some differences in the construction of the Ginis, such as from net vs. gross income and personal vs. household income. We tackle these problems by using first a $PERSON$ dummy in the regression where if the Gini observation is based on personal income.[30] Also, different countries may have different tax systems and inequality in gross incomes can be mitigated by certain tax policies. Hence we use a NET dummy where such a construction is reported. Finally, for the purposes of enlarging our data set, we include a few more Gini data to our data set that were classified as non-acceptable by D-S; however, we do control for this problem by using the NA dummy in the regression.[31] Another issue about the data set is that we construct it in the form of 5-year intervals.[32,33]

4.3 Estimation Issues

We first estimate the Equations (1) - (3) with Ordinary Least Squares (OLS), with and without control vectors of **X**, **Y** and **Z.** Simple OLS is useful to get a sense of the data and the model behavior, especially in the context of an empirical verification of theoretical predictions.

However, the econometric models specified in Equations (1) - (3) are the structural equations of a simultaenous equations system; this may require an instrumental variables (IV) estimation. A careful reader will notice that the joint determination of the variables in this model is recursive, that is, the system is triangular. The endogenous variables of the system are α, $\frac{(1-A)}{A}$ and $INEQ$, and there is a unidirectional dependency among the endogenous variables. The first equation is completely determined by exogenous factors, in particular k, and the second endogenous variable is determined by the first endogenous variable, and so on. Technically, this system can be estimated in a reduced form by forming a single inequality equation by forward substitution of the endogenous variables. But we subscribe to the estimation of separate equations in order to test the

---

Spilimbergo *et al.* (1999) report that a logarithmic transformation changes the sign of per capita income variable in the inequality regressions. Schultz (1998) finds contrasting signs, while Barro (2000) finds supporting signs. Li et al. (1998) report that Kuznets hypothesis finds support in the cross-country dimension of the data set, rather than within country variation over time.

[30] Kuznets revises his 1962 and 1974 papers in his book Economic Development, The Family and Income Distribution (1989), and argues that the recipient unit to be considered in internal inequality should be family or household, not person. See p. 132.

[31] For example, Barro (2000) adds some Gini data to his data set that was classified non-acceptable by D-S, and then uses a dummy variable in the regression to control for this action. But, in this practice, any measurement error in Gini should be captured by the error term, because Gini is the dependent variable.

[32] Most authors resorted to doing this in order to deal with the problems on the unbalanced Gini data (See Barro (2000) who does it with 10-year intervals.). Therefore we take the 5-year averages of data on every variable, including the Ginis, and assign them to the relevant intervals.

[33] Ginis are based on expenditures vs incomes in the D-S dataset. As suggested by D-S (1996), we add 6.6 Gini points to the expenditure-based Ginis to obtain a consistent series.

theoretical predictions, i.e., to trace the impact of the right-hand side variables on the endogenous variables within the mechanism offered.

For estimating the parameters of a triangular system (which are the elements of an upper triangular matrix), one needs to figure out first if the structural disturbance covariance matrix of the system is diagonal; i.e., whether there is contemporanous correlation across the equations or not. If there is such a correlation, one needs to secure the identification of the system first, and then proceed with simultaneous equations estimation methods (see Greene (2000, p. 393)). Otherwise, the system is said to be identified, and OLS would provide consistent results. To determine whether a system method is necessary and in that respect whether the joint estimation of these equations would provide efficiency gains owing to cross-equation information, we run a test suggested by Breusch and Pagan (1980). In particular, the squares of the correlations between the residuals of each equation are multiplied by the relevant sample size, and their summation provides a chi-squared test statistic with degrees of freedom equal to the number of equations. It turns out that the test result with the base sample provides some evidence for the possibilty of a non-diagonal covariance matrix, while the other samples do not have the indicia for such correlation.

Our first equation is already identified by construction. Identifying the second and the third equations necessitates meeting the rank and order conditions of the model. In doing this, we employ a general rule of thumb, in which case we secure that each equation has its own distinct exogenous variable. Indeed, all the controls are assumed to be exogenous.[34]

The endogeneity of k to the left-hand side variables should also be addressed. Although there are exogenous initial elements that increase k, once farmers start migrating and becoming workers in manufacturing, k is expected to increase in Equation (2). Reasonably, this situation should apply to all countries. k may additionally be endogenous to α and $INEQ$ due to reasons outside our model. Although our interest is more on the orientation (sign) of the coefficients rather than the magnidutes (which would be biased and inconsistent due to endogeneity), for a prudent econometric treatment, we assume that the problem exists and estimate the system with Generalized Method of Moments (GMM). Indeed we test for endogeneity formally in the Section 6 too. GMM is robust to heteroskedasticity of unknown form. As instruments, we use the exogenous variables of the models and the lagged values of the endogenous variables relevant in each equation (including that of k).

With all care taken in the econometric analysis, however, we do not find any systematic difference between the OLS and GMM results. Therefore in the discussions we will focus on the OLS results, as with GMM estimations we lose certain observations belonging to four countries of the base sample and many others in the other samples, due to having to use lagged values of the variables as instruments. From a broader perspective, the GMM estimation can be seen a robustness check of our basic OLS results with a different econometric methodology.

[34] With number of instruments greater than the number of parameters in our equations, the system is said to be over-identified; we conduct the suggested over-identification tests and the null is accepted in every case.

4.4 Control Variables

For vector $\mathbf{X}$ of Equation (1), we believe that people's tastes would be affected by demographic, geographic, and cultural factors.[35] The share of arable land in countries' total area can be used in a cross-country setting to control for the effects of agricultural land availability on the tastes of people. Likewise, high population density in rural areas may result in more affinity for agricultural goods. However, high rural population density may also increase on-farm consumption of the agricultural goods, and therefore reported tastes (through the agricultural GDP) may be underestimated. Which effect is stronger is an empirical question. Additionally, religious factors may be part of cultural effects to determine the affinity for agricultural goods. Finally, small island countries may have lower agricultural tastes; i.e, production of cash crops as against income that can be generated from tourism services.

Regarding the control variables in Equation (2), $\frac{(1-A)}{A}$ can be perceived in two ways: i) the relative manufacturing/agricultural labor force, and thus the capacity of manufacturing to that of agriculture, and ii) the accumulated migration per farmer. In the case of i), initial schooling can capture the potential of the economies to grow into manufacturing.[36] Also, it would be useful to control oil producing countries in the sample. In the case of ii), total surface area of the countries can indicate the spread of the settlements, division of urban and rural settlements within the country, and thereby mobility of the labor force. Being landlocked and the latitude may impact both i) and ii) in that while the former can affect the accessibility to capital goods from the world markets, the latter can point out to natural resource endowments and influence the efficiency and productivity in production.[37]

In Equation (3), it will be useful first to control the different levels of inequality across countries and special features in the construction of the Gini data. Schultz (1998), among others, proposed that regional differences explain the levels of inequality to an important extent; therefore, we use regional dummies in the regressions (Latin American and Caribbean countries, East Asian and Pacific countries etc).[38] We also use $NET, PERSON$ and NA dummies to control for differences in Gini construction. In addition, Rodrik (1999b) finds that democracies pay higher wages, so we control for the level of democracy in explaining the income distribution too. In the sensitivity analysis section, we use alternative control variables, as suggested by Li *et al.* (1998).

4.5 Other Data Sources

We provide the sources of other data in Appendix D. However, note that most variables in our estimations are in percentages, not in decimals, to make the interpretation easier. Summary statistics for the main variables are provided in Table 1.

[35] The economic factor is captured by k and its possible impact on national income.

[36] Papageorgiou (2003) finds that primary schooling contributes to the productive capacity of the economies, while post-primary schooling adds to their innovative capacity.

[37] For arguments on the influence of natural resources and latitude on the business environment, see Acemoglu *et al.* (2001) and Easterly and Levine (2003).

[38] The level differences in Gini could have been controlled by lagged Gini, had there been continuity in the country observations over time.

5 Results

5.1 OLS Results with No Controls

Table 2a presents the OLS results for the predicted specification with no controls. The purpose of this preliminary analysis is to look at the unconditional relationship between the state variables and the endogenous variables of our model.

In Equation (1), the impact of k on α is insignificant across the samples, although the sign is negative in the base sample. In Equation (2), the impact of α on $\frac{1-A}{A}$ is highly significant at 1% level, with the expected negative sign, implying that a decrease in agricultural tastes, hence the income of the farmers, leads to migration. The impact of k is estimated to be insignificant in the base and base + Sub-saharan African countries cases, while the samples that include developed countries deliver significant and (unexpected) negative sign for k - a result that can be attributed to developed countries only. As per Equation (3), α is significantly estimated at 1% level across the samples with the expected negative sign. $\frac{1-A}{A}$ is also significantly estimated across the samples with expected negative coefficients. Lastly, k is estimated to be insignificant across the samples, although it is marginally significant with the predicted positive sign in the all-countries case.

These preliminary results indicate that, if initiated somehow, the migration mechanism is associated with inverted-U shape in income inequality (i.e., through the negative and significant sign of α on $\frac{1-A}{A}$ and the negative and significant signs of α and $\frac{1-A}{A}$ on $INEQ$). Also, there is some evidence for its higher effectiveness in the developing countries. The insignificance of k is most likely related to the absence of controls and is addressed at various forthcoming parts of the paper.

5.2 OLS Results with Controls

The results so far assume that, apart from the state variables there are no other factors affecting the endogenous variables. This is unlikely because, at least empirically, there would be a number of factors at work in the determination of α, $\frac{(1-A)}{A}$ and $INEQ$. Therefore, controlling for some of the other factors should lead to a better and a more precise picture in our estimations.

The estimation results with control variables improve upon the preliminary results and provide noticeably significant and anticipated signs for the state variables. In Equation (1), k is estimated to have a highly significant (at 1% level) and negative effect on α in the base and base + developed countries cases (with coefficients around -0.040 for both), while the samples that include the Sub-saharan African countries provide no significant relationship. This implies that, holding other factors constant, Kuznets' observation on the negative relationship between α and k is working in developing (excluding Sub-saharan African) and developed countries, while the cited effect in so insignificant in the Sub-saharan African countries that it overrules that significant effect found in the other countries. The coefficient -0.040 implies that α falls by 4% with every additional 100 workers per firm in the economy, holding other factors constant.

In Equation (2), α is estimated to have a strongly significant (at 1% level) negative effect on $\frac{(1-A)}{A}$. The coefficients are significant across all samples, while their magnitudes range between -0.027 and -0.055. The base sample has the second highest impact with a coefficient -0.049. This implies that in the base countries every 20% decline in α (e.g. from 50% to 30%) is associated with a higher relative labor force ratio by 1 unit (for example, in terms of division of the labor force between non-agricultural and agricultural sectors, a change from 50%-50% to 67%-33%, or from 67%-33% to 75%-25%). In this equation, k is estimated to have a significant migration-fostering impact in the base and base+Sub-saharan African countries, while the samples with developed countries do not indicate such relationship. The estimated significant coefficients imply that every additional 200-300 workers per entrepreneur is associated with a higher relative labor force ratio by 1 unit, holding other factors constant.

In Equation (3), we estimate α with significant and negative coefficients across all samples. The magnitude of the coefficients are between -0.129 and -0.312 but the highest effect belongs to the base countries. As per $\frac{(1-A)}{A}$, in all cases is it estimated with significant and negative coefficients. This negative implies that the countries with higher migration experience would have lower inequality. The samples with base and base +Sub-saharan African countries provide significantly higher coefficients (around -1.3) than the samples with developed countries (which are around -0.6). This reaffirms the point above on the effectiveness of the migration-inequality process in different set of countries; i.e., higher effectiveness in developing countries. It also implies that the curvature of the inverted U-shape owing to this mechanism is steeper in those countries. Lastly, the data do not provide support for the direct impact of k on income Gini. This might be due to the strict literal test that we are undertaking and/or the data quality, and therefore we explore this issue further in the sensitivity analysis section by relaxing our approach.

Thus, overall, while the indirect effect of a change in k on the migration process is not working in Sub-saharan African countries (from Equation 1), the direct effect is not working in the developed countries (from Equation 2). On the other hand, both effects are working in the base countries. In addition, it is found that migration, as shown by an increase in $\frac{(1-A)}{A}$, has equilibriating effect on income distribution in all countries, although its effect is relatively more pronounced in the base and base+Sub-saharan African countries (from Equation 3).

There are also important implications in the estimation of the control variables. For instance, in Equation (3), a higher democratic score is associated with lower inequality, verifying Rodrik's argument on democracy-wages relationship. There is around 6-8 Gini points difference between the most free and the most dictatorial regime, holding other factors constant. In Equation (1), population density in rural areas has a significant positive sign, favoring the affinity argument. Interestingly, samples with Sub-saharan African countries provide higher coefficients. Small island dummy has a significant and negative sign again with varying coefficients across samples. In Equation (2), schooling in 1960 and oil producing countries are associated with higher level of relative labor force across the board, while higher surface area lowers the ratio.[39] Finally, although not shown in the tables regional dummies add relatively litte to the explanation of inequality in the base countries (see the Adjusted R-squared), probably because our state variables are able to catch those effects, while all controls add to the explanation of inequality in the other samples.

[39] The control dummies on the construction of the income gini never had any explanatory power in the regressions, except NA in a few cases, in which case all three are tested to be jointly equal to zero.

5.3 GMM Results

In the bottom panel of Table 2b, the Breusch-Pagan test results are presented. Among all samples, the base countries case provides the highest test statistic (with a p-value 0.21). We deem this as an evidence for cross-equation information. The resulting GMM estimations (presented in Table 3) show that all of our results remained qualitatively similar to the OLS estimations, while we obtain some efficiency gains for, in particular, the α equation.

5.4 A Total Effect Analysis

Our empirical analysis tests only the predictions of the theoretical model by looking at the signs of the coefficient estimates. The theoretical model is built with a number of assumptions and restrictions and as such does not predict any particular magnitude. Thus, it may be misleading to use the coefficient magnitudes to find the numerical relationship in the k-migration-inequality mechanism. However, it may be of interest to carry out such an exercise with GMM results to understand the mechanism: The difference between the lowest and the highest k in our base sample is 456.5 (the lowest belongs to Iran (65-69) with 3.1 and the highest belongs to Guyana (80-84) with 459.4). For such difference, our estimated model predicts an 21.46% (456.5×-0.068) lower α value for Guyana, holding other factors constant. In the sample, Iran has an α value of 37% and Guyana has 4%. Given other effects such as the size of the countries and other consumption habits of the people, the difference between the estimated value and the actual value is plausible. The higher k and the lower α create a potential for inter-sectoral income gap and cause migration. Thus the model predicts a higher $\frac{(1-A)}{A}$ for Guyana by 3.7 unit ($-0.068 \times -21.46 + 456.5 \times 0.005$). In reality, the difference is 1.8; Iran has a $\frac{(1-A)}{A}$ of 1.1 and Guyana has a value of 2.9 (the division of the labor force across sectors is roughly 53%-47% for Iran and 74%-26% for Guyana). Admittedly. k is only one of the reasons behind the fall in agricultural income and thus there should remain some uncaptured effects on migration. Hence, the difference between 1.8 and 3.7 is understandable. Note again that the origin of this difference goes to differing ks at the beginning. Finally, the 21.46% difference in α should lead to Guyana having a 6.8 higher Gini point (-21.46×-0.317, where the k effect does not work on Gini). The following migration effect, after having worked, should decrease Gini by 4.6 points (-1.247×3.7). Thus the model predicts still a 2.2 higher Gini points for Guyana, holding other factors constant. In reality, Guyana has a Gini of 56 (gross, household, acceptable) and Iran has a Gini of 49 (net, person, acceptable).[40] Noting again that our purpose is to confirm the theoretical signs with data only, we think that the estimations deliver plausible magnitudes for the k-migration-inequality mechanism. The differences should be reconciliable given all the measurement issues on Gini and many other determinants of the endogenous variables, as well as the theoretical and empirical restrictions on the analysis.

[40] Net income should be distributed more evenly than gross income due to progressive taxation systems, which implies that Iran would have a higher gross income Gini.

6 Sensitivity Analysis

We check the sensitivity of our results on a number of grounds.

Sensitivity of the regressors. In Equation (1), we tried various specification checks such as excluding the control variables one at a time and using their different combinations. It turns out that all controls are jointly important for the significance of k in the base and base + developed countries.[41]

In the equation of $\frac{1-A}{A}$, while all controls are jointly important in the base + Sub-saharan African sample, the oil dummy is somewhat more important for a significant k in the base sample. α remains robustly significant in all checks.

As per Equation (3), Li *et al.* (1998) explain international and intertemporal variation in inequality on the basis of political economy and capital market imperfection arguments. They find that civil liberties, initial level of schooling (of 1960), land Gini (i.e., asset inequality), and $M2/GDP$ (indicator of financial development and credit opportunities) are important determinants of inequality. We include all these variables into our specification (except civil liberties, which was already included as part of democracy), and it turns out that these variables have no explanatory power. However, α, $\frac{1-A}{A}$ and democracy remain strongly significant with their original signs, while k is still insignificant with this exercise. Removing the regional dummies from the regression does not change this result. But, removing our state variables from the regression makes land Gini positive and significant (with also democracy being significant), implying that our state variables can capture the political economy part of Li *et al.*'s argument. The insignificance of the financial development controls can be attributed to the closed economy period we are looking at.

One can argue that certain controls in $\mathbf{X}$, $\mathbf{Y}$ and $\mathbf{Z}$ may belong to the other equations as well. By retaining one distinct variable in each equation (i.e., log area in Equation (2) and democracy in Equation (3)), we re-estimate the equations with all controls included. Although we lose the significance of k in Equation (2) and that of α in Equation (3), we retain the other state variables significant. This is important to see both that the suggested mechanism works and the original choice of the controls for the respective equations is relevant.

Econometric concerns. We have noted that there are initial exogenous elements that change k in the system. However, once migration starts, the right-hand side variable k in Equation (2) is also expected to increase, suggesting an endogeneity. k may also be endogenous in Equations (1) and (3) due to reasons not endogenized in our model for tractability reasons. For instance, decreasing international vis-a-vis domestic prices may change agricultural output, influencing α, and switching output composition towards manufacturing, hence increasing k. Although openness is not a part of the theoretical model as well as our sample, a measure should be taken. Moreover, when the government is an important employer in the economy, hiring decisions can be based on redistribution concerns in an unequal society, suggesting an endogeneity for k in Equation (3). Thus we carry out a Hausman (1978) test, which tests whether the OLS and instrumental

[41] It is evident that per capita income in economies is an important and strong determinant of agricultural tastes (Engel-law idea). We do not include this variable in the α equation, which otherwise would strip out the explanatory power of k as an income-gap causing variable (it is easy to observe that our theory implies a strong relationship between k and per capita income, and theoretically and empirically per capita income should comprise the effects of k).

variables coefficients are significantly different than each other.[42] The testing procedure is reported in Table 4. The test results indicate that k is endogenous to $\frac{1-A}{A}$ in all samples except the base + Sub-Saharan African sample (in essence, the no-endogeneity effect in this sample is dominated by the Sub-Saharan African countries). There is also evidence for endogeneity of k in Equation (1) in the samples with Sub-Saharan African countries, while none in the others. Endogeneity in Equation (3) is found only in the base sample. Thus, instrumenting k with its lagged value and running a GMM estimation was an appropriate strategy from econometric point of view, even though our overall results did not change. The state variables α and $\frac{1-A}{A}$ remain strongly significant in the respective regressions with this exercise.

Additionally, we find evidence for the existence of heteroskedasticity of k in Equations (1) and (2) with White's test (in single-equation context). Therefore the use of GMM is justified from another aspect.[43]

Alternative Variables. Although the role of k in Equation (3) is not as critical as those in Equations (1) and (2) for the suggested mechanism to be initiated, its insignificance in this equation can be investigated.

It is shown that the effects of k on income inequality work through α and $\frac{1-A}{A}$ in a certain set of countries. However, the insignificance of the direct impact raises a question: empirically, what does the value of k refer to? Well, it may not be standard across countries. For instance, 100 workers of Bangladesh may produce the same product as 60 workers of Colombia, and if we look at only the Between-variation in k, Bangladesh should have higher inequality than Colombia, other things being equal.[44] On the other hand, an α value 50% implies twice as strong agricultural tastes than 25%, and similarly, the Gini coefficient 30 implies twice as better income distribution than 60 (if the measurement differences are controlled for). So it may be useful to look at only the Within-variation in k (i.e., change in k).

Table 5 reports the results where k is replaced with Δk in all three equations for the base sample. We report the results with OLS (no controls and with controls) and GMM. To start with Equation (3), Within-variation in k (in short, Within-k) increases Gini in all cases, the impact being significant at 10% in the OLS with no controls case; the significance weakens, but still plausible, with the addition of controls. Thus, this implies that it may be the 'extent' of a change in k that can affect the income distribution. With this exercise, α and $\frac{1-A}{A}$ remain strongly significant with the predicted signs in this equation. In Equation (1), Within-k has the predicted negative sign, significant in the OLS with no controls case and weakly significant in the other cases. This implies that although the change in k has some explanatory power for α, its Between-variation across countries is also important to find the cited effect. In Equation (2), Within-k is

[42] In an auxiliary regression, k is regressed on a set of exogenous variables and the residuals from this regression are then inserted in the original regressions. The significance of the residuals would indicate endogeneity. We employed various sets of exogenous variables in the auxiliary regressions and the evidence on endogeneity was varying. For consistency, we resorted to using only the lagged value of k in the auxiliary regressions, and the test results are based on this instrument.

[43] In estimating our models, one can also suggest to use a fixed-effects estimator. However, note that we address the country heterogeneity to a great extent by employing different sub-samples of relatively homogenous countries. Thus, we do not resort to such estimator.

[44] Let us remind the reader that we are undertaking a literal analysis and as such do not account for the role of capital per worker in production etc. both in theory and empirics. However, in reality the data generating process can differ due to various observed and unobserved reasons.

estimated to be insignificant in all cases, implying that the Between-variation in k is more important in explaining the relative labor force. α remains strongly significant with the predicted sign as a result of this exercise.

In an alternative system specification, we employ only the Within-variation of all state variables in the estimation (all time-invarying variables are eliminated in this specification); however, we do not obtain significant estimates with this experiment, pointing out the importance of the Between-country variation in results.

Also, one could argue that the empirical measure of k, the ratio of number of workers to the number of establishments in countries, is too literal to use in the empirical analysis. In Table 6, we check the relevance of k in representing the relative size of manufacturing and agricultural sectors across countries. One intuitive variable to check k against is the ratio of manufacturing to agricultural value added (in short, MV/AV). We report the GMM results for only the base sample.[45] Just as in the relationship between α and k, MV/AV has a negative and strongly significant effect on α. It also has the predicted positive sign in Equation (2), being strongly significant at 1%. The sign of α does not change with this replacement in Equation (2), however its significance diminishes. In Equation (3), MV/AV has the predicted positive sign, but estimated to be insignificant. This result is the same as the result with k. The Within-variation in MV/AV (unreported) does not deliver significant results. Nevertheless, both α and $\frac{1-A}{A}$, which are the key variables in the migration-inequality relationship, remain strongly significant with predicted signs in this equation. Thus, owing to similar results, we conclude that k represents effectively the relative size of manufacturing to agriculture sectors.[46]

7 Concluding Remarks

Our starting point in this paper is the widespread inconclusiveness on the empirical performance of Kuznets' hypothesis. We ascribe this outcome to imprecise domains in which the Kuznets hypothesis has been formulated and analyzed. Testing the inverted-U hypothesis only in terms of its shape, virtually independent of Kuznets' own reasoning, contributed to the inconclusiveness further.

According to Kuznets (1973), p. 255), prominent features of the pre-industrial phase have been "a low per capita product, a large share of agriculture and other extractive industries, a generally small scale of production" which are captured by our modeling of that sector. The transition out of the pre-industrial phase entails "[t]he sustained rise in the supply of goods," ("[a]dvancing technology [as] the permissive source" of it) as well as "a shift away from agriculture to nonagricultural pursuits" (Kuznets (1973), pp. 247-8). In our study we take the level of k to indicate the intersectoral firm size ratio as well as the level of technological advancement in manufacturing. An insight of Kuznets also allows us to incorporate a negative link between k and α. Thus, in this paper, instead of taking the above process simply as manifestation of a negative link between per capita income and inequality, in the light of Kuznets' insights we depict that process as one which links k, α and migration.

[45] Endogeneity and the Breusch-Pagan tests (unreported) show that MV/AV is endogenous to the dependent variables in all equations. The p-value for the Breusch-Pagan test is 0.15.

[46] We also regress MV/AV on k along with the other controls of the respective equations (i.e., Equations (1), (2) and (3)), and in the first two equations, we obtain positive and significant relationship.

We later test the theoretical predictions of the model. Taking the rising trend in k within countries as the starting point, we investigated empirically the indirect and direct effects of change in k on migration-inequality link across different samples. The data provide relatively strong and robust support for the suggested relationships with expected signs and high significance levels in relatively homogenous lot of countries, while the support decreases when other set of countries are used, as implied by the Kuznets' reasoning.

In this paper we focused only on the phase of industrialization at which developing countries relied heavily on import substitution policies. During that phase, clearly the role of migration as an equilibriating mechanism can not be overstated. In the last few decades, however, many countries switched from import substituting (IS) regimes to export promoting (EP) ones. This switch did not seem to slow down migration much in most of these countries. It would be interesting to see whether the inequality has increased or decreased in these countries following the regime switch and what kind of processes accounted for changes in inequality.[47]

References

[1] Acemoglu, D., J. A. Robinson and S. Johnson (2001), "The Colonial Origins of Comparative Development: An Empirical Investigation", American Economic Review, 91, pp. 1369-1401.

[2] Acemoglu, D. and S. Robinson (2002), "The Political Economy of Kuznets Curve", Review of Development Economics, 6, 183-203.

[3] Aghion, P. and P. Bolton, (1997), "A Theory of Trickle-Down Growth and Development", Review of Economic Studies, 64, 151-172.

[4] Amsden, A. and R. van der Hoeven (1996), "Manufacturing Output, Employment and Real Wages in the 1980s: Labour's Loss until the Century's End," Journal of Development Studies, 32, 506-530.

[5] Anbarci, N. and M. A. Ulubasoglu (2005), "Endogenous Liberalization and Within-Country Inequality", mimeo, Florida International University and Deakin University.

[6] Barro, R. (2000), "Inequality and Growth in a Panel of Countries", Journal of Economic Growth, 5, 5-32.

[7] Bhatia, K. (1979), "Rural-Urban Migration and Surplus Labour," Oxford Economic Papers, 31, 403-414.

[8] Bollen, K. (1990), "Political Democracy: Conceptual and Measurement Traps", Studies in Comparative International Development 25:7-24.

[47] In a companion paper in progress, we first try to identify theoretically and empirically the circumstances under which different segments in the society (entrepreneurs, workers and farmers) would support a switch from the Import- Substitution (IS) regime to the Export-Promotion (EP) regime. Second, we try to identify (again theoretically and empirically) the circumstances under which the inequality will increase or decrease after a switch from from the IS regime to the EP regime.

[9] Breusch, T.S. and A.R. Pagan (1980), "Lagrange Multiplier Test and its Application to Model Specification in Econometrics", Review of Economic Studies, 47, 239-53.

[10] Bruton, H. (1988), "Import Substitution", Chapter 30 in Handbook of Development Economics, Vol. 2, (ed.s H. Chenery and T. N. Srinivasan, North-Holland: Amsterdam), 1601-1644.

[11] Calvo, G. (1978), "Urban Unemployment and Wage Determination in LDCs: Trade Unions in the Harris-Todaro Model," International Economic Review, 19, 65-81.

[12] Cole, W. E. and R. D. Sanders (1985), "Internal Migration and Urban Employment in the Third World," American Economic Review, 75, 481-494.

[13] Deininger, K. and L. Squire (1996), "A New Data Set Measuring Income Inequality", World Bank Economic Review 10(3): 565-91.

[14] Deininger, K. and P. Olinto (1999), "Asset Distribution, Inequality and Growth", World Bank Policy Research Paper Series No. 2375.

[15] Easterly, W. and R. Levine (2003), "Tropics, Germs, and Crops: the Role of Endowments in Economic Development", Journal of Monetary Economics, 50:1, 3-39.

[16] Gang, I. and S. Gangopadhyay (1985), "A Note on Optimal Policies in Dual Economies," Quarterly Journal of Economics, 100, 1067-1071.

[17] Gang, I. and S. Gangopadhyay (1987), "Optimal Policies in a Dual Economy with Open Unemployment and Surplus Labour," Oxford Economic Papers, 39, 378-387.

[18] Greene, W. (2000), "Econometric Analysis", Prentice Hall.

[19] Hausman, J. (1978), "Specification Tests in Econometrics", Econometrica, 46, 1251-71.

[20] Harris, J. R. and M. P. Todaro (1970), "Migration, Unemployment and Development: A Two-Sector Analysis," American Economic Review, 60, 126-142.

[21] Hoddinott, J. (1999), "Wages and Unemployment in an Urban African Labour Market," Economic Journal, 106, 1610-1626.85),

[22] Kuznets, S. (1955), "Economic Growth and Income Inequality," American Economic Review, 45, 1-28.

[23] Kuznets, S. (1973), "Modern Economic Growth: Findings and Reflections," American Economic Review, 63, 247-258.

[24] Kuznets, S. (1966), "Modern Economic Growth: Rate, Structure, and Spead", Yale University Press: New Haven, CT.

[25] Kuznets, S. (1989), "Economic Development, the Familty and Income Distribution: Selected Essays", Cambridge University Press: Cambridge, England.

[26] Larson, D. and Y. Mundlak (1997), "On the Intersectoral Migration of Agricultural Labor," Economic Development and Cultural Change, 45, 295-319.

[27] Levy, V. and J. L. Newman (1989), "Wage Rigidity: Micro and Macro Evidence on Labor Market Adjustment in the Modern Sector," World Bank Economic Review, 3, 97-117.

[28] Lewis, W. A. (1954), "Economic Development with Unlimited Supplies of Labor," Manchester School of Economic and Social Studies, 22, 139-191.

[29] Li, H., L. Squire and H. Zou (1998), "Explaining International and Intertemporal Variations in Income Inequality" Economic Journal, 108(446): 26-43

[30] Lindert P. H. (1986), "Unequal English Wealth since 1670," Journal of Political Economy, 94, 1127-1162.

[31] McMillan, J. and A. Barkley, 1992, "Political Freedom and the Response to Economic Incentives: Labor Migration in Africa, 1972-87.", Institute for Policy Reform Working Paper Series. Washington D.C.: Institute for Policy Reform.

[32] Otsuka, K., H. Chuma and Y. Hayami (1992), "Land and Labor Contracts in Agrarian Economies: Theories and Facts," Journal of Economic Literature, 30, 1965-2008.

[33] Papageorgiou, C. (2003), "Distinguishing Between the Effects of Primary and Post-Primary Education on Economic Growth", Review of Development Economics, 7:4, pp. 622-635.

[34] Quibria, M.G. (1988), "Migration, Trade Unions and the Informal Sector: A Note on Calvo," International Economic Review, 29, 557-563.

[35] Ranis, G. and J. Fei (1961), "A Theory of Economic Development," American Economic Review, 51, 533-556.

[36] Ray, D. (1998), "Development Economics", Princeton University Press: Princeton, NJ.

[37] Rodriguez, F. and D. Rodrik (2000), "Trade Policy and Economic Growth: A Skeptic's Guide to the Cross-National Evidence," NBER Macroeconomics Annual, 15, 261-325. Cambridge and London: MIT Press.

[38] Rodrik, D. (1998), "Globalization, Social Conflict and Economic Growth," The World Economy, 21, 143-158.

[39] Rodrik. D. (1999a), "The New Global Economy and Developing Countries: Making Openness Work. Policy Essay No. 24. Overseas Development Council: Washington, D.C.

[40] Rodrik, D. (1999b) "Domocracies pay higher wages," Quarterly Journal of Economics, 114, 707-738.

[41] Rodrik, D. (2001), "The Global Governance of Trade as If Development Really Mattered," mimeo, Kennedy School of Government, Harvard University (prepared for UNDP).

[42] Sachs, J. and A. Warner (1995), "Economic Reform and the Process of Global Integration", Brookings Papers on Economic Activity, 0(1): 1-95.

[43] Schultz, T. P. (1998), "Inequality in the Distribution of Personal Income in the World: How It Is Changing and Why", Journal of Population Economics, 11, 307-44.

[44] Sjaastad, L. A. (1962), "The Costs and Returns of Human Migration," Journal of Political Economy, 70, 80-93.

[45] Smith, A. (1937), "The Wealth of nations, The Modern Library: New York, NY.

[46] Spilimbergo, A., J. Londono and M. Szekely (1999) "Income Distribution, Factor Endowments, and Trade Openness", Journal of Development Economics, 59(1): 77-101

[47] Stiglitz, J. (1974), "Alternative Theories of Wage Determination and Unemployment in LDCs: The Labor Turnover Model," Quarterly Journal of Economics, 88, 194-227.

[48] Timmer, C. P. (1988), "The Agricultural Transformation," Chapter 8 in Handbook of Development Economics, Vol. 1, (ed.s H. Chenery and T. N. Srinivasan, North-Holland: Amsterdam), 276-329.

[49] Todaro, M. (1969), "A Model of Labor Migration and Urban Unemployment in Less Developed Countries," American Economic Review, 59, 138-148.

[50] Tybout, J. R. (2000), "Manufacturing Firms in Developing Countries: How Well Do They Do, and Why?" Journal of Economic Literature, 38, 11-44.

[51] United Nations, UNIDO Industrial Statistics Database, 2003.

[52] Wacziarg, R. and K.H. Welch (2003), "Trade Liberalization and Growth: New Evidence", NBER Working Paper No. 10152.

[53] Willamson J. G. (1985), "Did British Capitalism Breed Inequality?", Allen and Unwin: Boston, MA.

[54] World Bank, World Development Indicators, 1999, 2003.

A Proofs of the Theoretical Results

Proof of Proposition 1: (1) A skilled worker's utility maximization problem can be rewritten as $u_w = L_w c_w$,where $c_w = a^{\alpha^m} m^{1-\alpha^m}$stands for the composite consumption good, such that $c_w = w_m(1 - L_w)$.

First order conditions from this problem yield $w_m L_w = c_w$. Plugging c_w into the constraint we get

$$L_w^* = \frac{1}{2} \Rightarrow l_w^s = \frac{1}{2} \tag{4}$$

The entrepreneur's maximization problem can be rewritten as $u_e = L_e c_e$,where $c_e = a^{\alpha^m} m^{1-\alpha^m}$stands for the composite consumption good, such that $c_e = w_m(1 - L_e) + \pi_m$.

First order conditions from this problem yield $w_e L_e = c_e$. Plugging c_e into the constraint we get $L_e = \frac{1}{2}(\pi_m + w_m)/w_m = \frac{1}{2}(\frac{\pi_m}{w_m} + 1)$

Thus, if $\frac{\pi_m}{w_m} \geq 1$, we will have $L_e^* = 1$ (since L_e cannot exceed 1). Suppose that the entrepreneur believes that $\frac{\pi_m}{w_m} \geq 1$ will take place and thus he/she chooses not to work. We will show that in equilibrium the entrepreneur's belief and action will be justified.

First-order conditions of π_m yield $k(\frac{1}{2}(l_m)^{-\frac{1}{2}} - w_m) = 0$. Thus, demand for labor per entrepreneur is $l_m^d = k(\frac{1}{2w_m})^2$. But $l_w^s = \frac{1}{2}$ (and $l_e^s = 0$). Thus, equating l_m^d and the total labor supply in the firm kl_w^s, the equilibrium manufacturing wage becomes

$$w_m = (\frac{1}{2})^{\frac{1}{2}} \tag{5}$$

Thus, by using (1.1) and (1.2), we get

$$\pi_m^* = k(\frac{1}{2})^{\frac{1}{2}} - k\frac{1}{2}(\frac{1}{2})^{\frac{1}{2}} = k\frac{1}{2}(\frac{1}{2})^{\frac{1}{2}} \tag{1.3}$$

Since $\pi_m^* = k\frac{1}{2}(\frac{1}{2})^{\frac{1}{2}} > (\frac{1}{2})^{\frac{1}{2}} = w_m$ and because k is an integer greater than 1, $L_e^* = 1$ and thus $l_e^s = 0$.

(2) By using (1.1) and (1.2) we get,

$$l_w w_m = \frac{1}{2}(\frac{1}{2})^{\frac{1}{2}} \tag{1.4}$$

(3) Let π_a^* denote the farmer's profit and w_a denote the farmer's wage. Observe that

$$Y_a^* = \pi_a^* + w_a l_a = p_a(l_a)^{\frac{1}{2}} \tag{1.5}$$

since $\pi_a^* = p_a(l_a)^{\frac{1}{2}} - w_a l_a$.

A farmer's utility maximization problem can be rewritten as $u_a = L_a c_a$,where $c_a = a^{\alpha^f} m^{1-\alpha^f}$stands for the composite consumption good, such that $c_a = w_a(1 - L_a)$.

First order conditions from this problem yield $w_a L_a = c_a$. Plugging c_a into the constraint and in Y_a^* we get

$$L_a^* = \frac{1}{2}, \quad l_a^s = \frac{1}{2}, Y_a^* = p_a(\frac{1}{2})^{\frac{1}{2}} \tag{1.6}$$

Thus, each farmer supplies

$$a^s = (\frac{1}{2})^{\frac{1}{2}} \tag{1.7}$$

First order conditions from any agent i's utility maximization problem yield

$$a_i = \frac{\alpha_i Y_i^*}{p_a} \tag{1.8}$$

where $\alpha_i = \alpha^f$ when agent i is a farmer and $\alpha_i = \alpha^m = 1/k$ when agent i is a worker or an entrepreneur.

Since there are two sectors, it follows from Walras' law that, if one of the markets is in equilibrium, then so is the other one. Using (1.6) and (1.8), the farmers' total demand for the agricultural good becomes

$$A\frac{\alpha^f Y_a^*}{p_a} = A\alpha^f(\frac{1}{2})^{\frac{1}{2}} \tag{1.9}$$

Using (1.4) and (1.8), the workers' total demand for the agricultural good becomes $\frac{k(1-A)}{k+1}\frac{\alpha^m Y_w^*}{p_a} = \frac{k(1-A)}{k+1}\frac{\alpha^m \frac{1}{2}(\frac{1}{2})^{\frac{1}{2}}}{p_a}$ but since $\alpha^m = 1/k$,the workers' total demand for the agricultural good becomes

$$\frac{(1-A)}{k+1}\frac{\frac{1}{2}(\frac{1}{2})^{\frac{1}{2}}}{p_a} \tag{1.10}$$

Using, (1.3) and (1.8), the entrepreneurs' demand for the agricultural good becomes $\frac{(1-A)}{k+1}\frac{\alpha^m Y_e^*}{p_a} = \frac{(1-A)}{k+1}\frac{\alpha^m k\frac{1}{2}(\frac{1}{2})^{\frac{1}{2}}}{p_a}$, but since $\alpha^m = 1/k$,the entrepreneurs' total demand for the agricultural good becomes

$$\frac{(1-A)}{k+1}\frac{\frac{1}{2}(\frac{1}{2})^{\frac{1}{2}}}{p_a} \tag{1.11}$$

Then, by using (1.9), (1.10) and (1.11), total demand for the agricultural good becomes

$$A\alpha(\frac{1}{2})^{\frac{1}{2}} + \frac{(1-A)}{k+1}\frac{(\frac{1}{2})^{\frac{1}{2}}}{p_a} \tag{1.12}$$

Using (1.7), the total supply of the agricultural good becomes

$$A(\frac{1}{2})^{\frac{1}{2}} \tag{1.13}$$

Then, using (1.12) and (1.13) (i.e., equating the total demand and supply of the agricultural good) yields $\frac{(1-A)}{k+1}\frac{(\frac{1}{2})^{\frac{1}{2}}}{p_a} = A(1-\alpha)(\frac{1}{2})^{\frac{1}{2}}$ where $\alpha = \alpha^f$.

By simplifying it, we get

$$p_a = \frac{(1-A)}{A}\frac{1}{(1-\alpha)(k+1)} \tag{1.14}$$

Then, by using (1.7) and (1.14)

$$Y_a^* = \frac{(1-A)}{A} \frac{1}{(1-\alpha)(k+1)} (\frac{1}{2})^{\frac{1}{2}} \tag{1.15}$$

This completes proof of Proposition 1.

Proof of Theorem 1: (1) We will calculate the individuals' indirect utility functions.

Recall by the proof of Proposition 1 that the leisure of the farmer and the worker is $\frac{1}{2}$ and that of the entrepreneur is 1. The first order conditions of u_a and u_w yield $a_a = \frac{\alpha^f Y_a^*}{p_a}, a_w = \frac{\alpha^m Y_w^*}{p_a}, m_a = \alpha^f Y_a^*, m_w = \alpha^m Y_w^*$, where $\alpha^m = 1/k$.

Thus, using the above first-order conditions and the income expressions in Proposition 1, the indirect utility functions become

$$u_a^* = \frac{1}{2}(\alpha^f)^{\alpha^f}(1-\alpha^f)^{1-\alpha^f}(\frac{1}{p_a})^{1-\alpha^f} Y_a^* \tag{T.1}$$

$$u_w^* = \frac{1}{2}(\alpha^m)^{\alpha^m}(1-\alpha^m)^{1-\alpha^m}(\frac{1}{p_a})^{1-\alpha^m} Y_w^* \tag{T.2}$$

where $\alpha^m = 1/k$.

Thus, a farmer will choose to migrate if $u_w^* > u_a^*$, and will be indifferent between migrating and not migrating if $u_w^* = u_a^*$,and will not choose to migrate if $u_a^* > u_w^*$. Thus, the equilibrium level of $\frac{A^*}{1-A^*}$ can be obtained when $u_w^* = u_a^*$. By (1.4) in the proof of Proposition 1, $Y_w^* = \frac{1}{2}(\frac{1}{2})^{\frac{1}{2}}$ and by (1.15) in that proof, $Y_a^* = \frac{(1-A)}{A}\frac{1}{(1-\alpha)(k+1)}(\frac{1}{2})^{\frac{1}{2}}$. Using some lengthy but straightforward algebra, $u_w^* = u_a^*$ yields

$$\frac{1-A^*}{A^*} = \frac{(1-\alpha^f)(k+1)[(\alpha^f)^{\alpha^f-\alpha^m} - (1-\alpha^f)^{\alpha^f-\alpha^m}]^{\frac{1}{-\alpha^f+\alpha^m}}}{2} \tag{T.3}$$

(2) In the proof of Part (1) of this theorem, when α^f and α^m are arbitrarily close, observe that by (T.3) $u_a^* \gtreqless u_w^*$ will reduce to $\frac{1-A^*}{A^*} \gtreqless \frac{(1-\alpha)(k+1)}{2}$.

Thus, when α^f and α^m tend to are arbitrarily close (in which case, for simplicity one can use $\alpha = \alpha^f = \alpha^m$), in equilibrium, we will have $\frac{1-A^*}{A^*} = \frac{(1-\alpha)(k+1)}{2}$.

Thus, $1 - A^*$ decreases in α^f (and thus α^m) and increases in k. This completes the proof of Theorem 1.

Proof of Theorem 2: By (1.3), (1.4) and (1.15) in the proof of Proposition 1 we have $Y_w^* = k\frac{1}{2}(\frac{1}{2})^{\frac{1}{2}}$, $Y_w^* = \frac{1}{2}(\frac{1}{2})^{\frac{1}{2}}$ and $Y_a^* = \frac{(1-A)}{A}\frac{1}{(1-\alpha)(k+1)}(\frac{1}{2})^{\frac{1}{2}}$.

Let s_i denote segment i's population share. The population weights of farmers, workers and entrepreneurs are $s_a = A, s_w = \frac{k(1-A)}{k+1}, s_e = \frac{(1-A)}{k+1}$ respectively. Then the average income Y^* becomes $\frac{(1-A)(k+\frac{1}{1-\alpha})}{(1-\alpha)(k+1)}(\frac{1}{2})^{\frac{1}{2}}$.

Let $t_i = \frac{Y_i^*}{Y^*}$. Thus, $t_e = \frac{1}{2}\frac{(1-\alpha)k(k+1)}{(1-\alpha)k(k+1)+1}$, $t_w = \frac{1}{2}\frac{(1-\alpha)(k+1)}{(1-\alpha)k(k+1)+1}$, $t_a = \frac{1}{2}\frac{A(k+1)}{(1-\alpha)k(k+1)+1}$.

Given these definitions, Gini coefficient can be calculated as follows (see equation (5) on p. 888 of Mookherjee and Shorrocks (1988)): $G = \frac{1}{2}[\sum_h \sum_k \mid t_h - t_k \mid]$.

Thus, we have $G = \frac{1}{2}(1-\alpha)\frac{(1-A)\frac{k(k-1)}{2(k+1)}+Ak-\frac{(1-A)}{(1-\alpha)}}{k(1-\alpha)+1}$.

Straightforward but lengthy calculations establish that G decreases in α, but increases in k and A. This completes the proof of Theorem 2.

B Other Data Sources and Definitions

The share of agricultural and manufacturing value added in GDP, arable land per capita, the share of arable land in total area, total surface area, labor force in agriculture, M2/GDP, rural population density and inflation data are obtained from World Development Indicators (1999, 2003). The data are the averages of each 5-year interval. Data on agricultural imports and imports are obtained from Food and Agricultural Organization (FAO) Trade Yearbooks (various issues starting from 1960). Schooling data for 1960 are obtained from Barro and Lee (2001). Land Gini data (obtained from Deininger and Olinto); religion variables (the shares of people affiliated to a certain religion in the total population -obtained from La Porta *et al.* (1999); small island, oil producing, and landlocked country dummies, latitude (absolute value of latitude is used in the regressions); and regional dummies (obtained from World Bank's compilation of Social and Fixed Factors) are time-invaryimg variables. Democracy scores for 1970-onwards (political rights and civil liberties averaged) are obtained from www.freedomhouse.org. Those for 1960 and 1965 are from Bollen (1990). For 1970 onwards, we used the averages of the scores for each interval (the data are available in 1-7 scale, where 1 is the most democratic. The data have been converted to 0-1 scale, where 1 is for the most democratic countries). Bollen data are available in 0-1 scale already, for the initial years of the 5-year intervals.

C Estimation Results

Table 1: Summary Statistics

Sample	Variable	Mean	Median	Max	Min	Std.Dev.	# Obs.	# Country
Base	k	76.60	57.77	459.44	3.11	63.98	81	32
	α (%)	21.36	19.20	63.98	3.53	13.12	82	32
	$\frac{1-A}{A}$	1.59	1.14	8.133	0.09	1.57	81	32
	$Gini$	45.86	47.54	62.00	30.06	7.99	78	32
	NET	0.32	0	1	0	0.47		
	$PERSON$	0.47	0	1	0	0.50		
	NA	0.12	0	1	0	0.33		
Developed	k	39.09	28.49	134.43	13.33	30.96	15	12
	α (%)	10.81	10.00	19.70	3.16	5.27	15	12
	$\frac{1-A}{A}$	6.82	7.37	16.67	0.99	4.59	15	12
	$Gini$	37.39	36.23	52.20	28.13	6.91	15	12
	NET	0.27	0	1	0	0.46		
	$PERSON$	0.27	0	1	0	0.46		
	NA	0.33.	0	1	0	0.49		
SubSaharan African	k	120.82	101.22	299.74	62.85	57.10	21	13
	α (%)	31.22	25.81	75.21	4.93	20.08	21	13
	$\frac{1-A}{A}$	0.96	0.30	5.85	0.09	1.58	21	13
	$Gini$	50.61	49.50	64.75	33.00	9.30	19	13
	NET	0.63	1	1	0	0.50		
	$PERSON$	0.84	1	1	0	0.38		
	NA	0.32	0	1	0	0.48		

Table 2a: Predicted Specification, No controls, OLS Estimations

Dep. Var.	**RHS Variables**	Base Sample	Base + Dev'ed.	Base + SS Afr.	Base+Dev'd +SS Afr.
	Constant	22.117***	19.409***	21.717***	19.087***
		(8.717)	(8.482)	(7.336)	(7.107)
α	k	-0.012	0.002	0.018	0.032
		(-0.509)	(0.072)	(0.584)	(1.022)
	Adj. R^2	0.01	0.01	0.01	0.01
	# Obs. (# Country)	80 (32)	95 (44)	101 (45)	116 (57)
	Constant	3.208***	5.283***	2.964***	4.659***
		(7.447)	(7.212)	(8.106)	(7.591)
	α	-0.074***	-0.122***	-0.060***	-0.092***
$\frac{1-A}{A}$		(-5.643)	(-6.166)	(-6.142)	(-6.591)
	k	-0.001	-0.007**	-0.001	-0.006**
		(-0.325)	(-2.239)	(-0.858)	(-2.369)
	Adj. R^2	0.36	0.27	0.13	0.26
	# Obs. (# Country)	80 (32)	95 (44)	101 (45)	116 (57)
INEQ	Constant	57.745***	51.725***	52.122***	48.517***
		(18.778)	(16.378)	(17.473)	(17.598)
	α	-0.405***	-0.265***	-0.229***	-0.163**
		(-4.183)	(-2.657)	(-2.582)	(-2.026)
	k	-0.011	0.002	0.006	0.015
		(-0.796)	(0.172)	(0.622)	(1.512)
	$(1-A)/A$	-1.300***	-0.945*	-1.016**	-0.822*
		(-2.699)	(-1.988)	(-2.063)	(-1.880)
	NET, PERSON, NA	Yes	Yes	Yes	Yes
	Adj. R^2	0.28	0.15	0.13	0.12
	# Obs. (# Country)	75 (32)	90 (44)	94 (45)	109 (57)

Notes: t-statistics in parentheses. *** means 1%, ** means 5%, and * means 10% significance. White heteroskedastic-covariance estimation.

Table 2b: Predicted Specification, with controls, OLS Estimations

Dep. Var.	RHS Variables	Base Sample	Base + Dev'ed.	Base + SS Afr.	Base+Dev'd +SS Afr.
	Constant	34.581***	3.406	1.397	-6.355
		(3.386)	(0.533)	(0.087)	(-0.831)
	k	-0.040***	-0.036***	-0.011	-0.003
		(-2.706)	(-3.044)	(-0.586)	(-0.122)
	Arable land	0.120**	0.093*	0.101	0.0702
		(2.197)	(1.817)	(1.171)	(0.901)
α	Log Rural PopDen.	2.807**	4.184***	6.042***	6.087***
		(2.461)	(5.542)	(3.513)	(5.218)
	Small Island	-10.211***	-11.621***	-18.071***	-17.198***
		(-2.748)	(-4.017)	(-3.611)	(-4.867)
	Religion Variables	Yes	Yes	Yes	Yes
	Adj. R^2	0.62	0.58	0.36	0.58
	# Obs. (# Country)	80 (32)	95 (44)	101 (45)	116 (57)
	Constant	3.691***	4.947	2.240**	3.332
		(4.694)	(1.971)	(2.364)	(1.662)
	α	-0.049***	-0.055***	-0.032***	-0.027***
		(-6.887)	(-3.774)	(-3.688)	(-2.539)
	k	0.003**	-0.000	0.002**	0.001
		(2.444)	(-0.069)	(2.029)	(0.650)
	Schooling 1960	0.206**	0.658***	0.400***	0.731***
		(2.474)	(8.452)	(4.538)	(8.787)
	Log Surface Area	-0.186***	-0.389*	-0.128*	-0.281*
$\frac{1-A}{A}$		(-2.528)	(-1.934)	(-1.659)	(-1.721)
	Landlocked	1.124***	1.282	0.023	0.078
		(2.591)	(1.533)	(0.083)	(0.236)
	Oil Prod.	2.966***	2.691***	2.634***	2.561***
		(5.240)	(4.944)	(5.057)	(5.775)
	Latitude	0.003	0.040**	0.012	0.023*
		(0.278)	(2.368)	(1.444)	(1.932)
	Adj. R^2	0.75	0.52	0.65	0.52
	# Obs. (# Country)	80 (32)	95 (44)	101 (45)	116 (57)
	Constant	57.843***	50.928***	55.669***	50.991***
		(15.269)	(11.156)	(15.209)	(13.129)
	α	-0.312***	-0.163*	-0.227***	-0.129*
		(-3.016)	(-1.743)	(-2.564)	(-1.637)
	k	-0.013	-0.007	0.002	0.007
INEQ		(-1.070)	(-0.639)	(0.275)	(0.724)
	$(1-A)/A$	-1.319***	-0.627*	-1.379***	-0.646*
		(-2.916)	(-1.893)	(-3.335)	(-1.847)
	Democracy	-7.331**	-8.328**	-5.916*	-8.187**
		(-2.086)	(-2.147)	(-1.710)	(-2.435)
	Regional Dummies	Yes	Yes	Yes	Yes
	NET, PERSON, NA	Yes	Yes	Yes	Yes
	Adj. R^2	0.31	0.36	0.21	0.27
	# Obs. (# Country)	74 (32)	89 (44)	92 (45)	107 (57)
BP test	χ^2 ~r	4.64	1.18	3.94	3.44
(1980)	p-value	0.21	0.76	0.28	0.35

Notes: See the notes to above tables.

Table 3: Predicted Specification, System (GMM) Estimations

Dep. Var.	RHS Variables	Base Sample	Base + Dev'ed.	Base + SS Afr.	Base+Dev'd +SS Afr.
	Constant	24.288	-2.746	-21.958	-16.168**
		(1.221)	(-0.653)	(-1.183)	(-2.392)
	k	-0.047***	-0.042***	0.027	0.028
		(-3.263)	(-2.914)	(0.557)	(0.665)
	Arable land	0.088*	0.090*	0.047	0.064
		(1.738)	(1.814)	(0.502)	(0.749)
α	Log Rural PopDen.	3.319**	3.929***	7.382***	6.251***
		(2.299)	(5.764)	(4.247)	(4.549)
	Small Island	-10.351***	-11.526***	-18.406***	-17.402***
		(-3.028)	(-4.528)	(-4.100)	(-5.116)
	Religion Variables	Yes	Yes	Yes	Yes
	Adj. R^2	0.60	0.60	0.30	0.34
	# Obs. (# Country)	64 (28)	75 (37)	82 (39)	93 (47)
	Constant	4.381***	3.889	2.411**	2.344
		(5.248)	(1.541)	(2.421)	(1.174)
	α	-0.068***	-0.062***	-0.037***	-0.028**
		(-5.391)	(-3.527)	(-3.288)	(-2.378)
	k	0.005*	-0.005	0.004**	0.005**
		(1.857)	(1.536)	(2.053)	(2.018)
	Schooling 1960	0.156*	0.625***	0.391***	0.705***
		(1.665)	(7.024)	(4.110)	(7.754)
	Log Surface Area	-0.219**	-0.300	-0.135	-0.202
$\frac{1-A}{A}$		(-2.450)	(-1.433)	(-1.543)	(-1.188)
	Landlocked	1.535**	1.304	-0.018	-0.044
		(2.338)	(1.177)	(-0.054)	(-0.128)
	Oil Prod.	2.897***	2.784***	2.714***	2.675***
		(5.587)	(5.630)	(5.209)	(5.820)
	Latitude	0.004	0.027	0.009	0.011
		(0.353)	(1.596)	(0.905)	(1.030)
	Adj. R^2	0.74	0.51	0.64	0.52
	# Obs. (# Country)	64 (28)	75 (37)	82 (39)	93 (47)
	Constant	59.179***	58.580***	57.065***	56.951***
		(10.631)	(10.597)	(13.207)	(13.867)
	α	-0.317***	-0.264**	-0.259***	-0.225**
		(-2.776)	(-2.369)	(-2.726)	(-2.505)
	k	-0.021	-0.022	-0.004	-0.006
INEQ		(-0.963)	(-1.033)	(-0.218)	(-0.360)
	$(1-A)/A$	-1.247***	-1.059***	-1.400***	-1.013**
		(-2.996)	(-2.653)	(-3.017)	(-2.427)
	Democracy	-6.407*	-8.251**	-4.841	-6.763*
		(-1.773)	(-2.179)	(-1.281)	(-1.813)
	Regional Dummies	Yes	Yes	Yes	Yes
	NET, PERSON, NA	Yes	Yes	Yes	Yes
	Adj. R^2	0.23	0.29	0.14	0.27
	# Obs. (# Country)	60 (27)	64 (31)	75 (38)	79 (42)

Notes: See the notes to above tables.

Table 4: Hausman Test for k

Dep. Var.	RHS Variables	Base Sample	Base + Dev'ed.	Base + SS Afr.	Base+Dev'd +SS Afr.
α	Constant	22.438	-2.814	-27.581	-18.446***
		(0.967)	(-0.624)	(-1.449)	(-2.663)
	k	-0.048***	-0.043***	0.032	0.029
		(-2.830)	(-2.530)	(0.928)	(1.062)
	Arable land	0.094*	0.094*	0.014	0.028
		(1.845)	(1.778)	(0.143)	(0.316)
	Log Rural PopDen.	3.520***	3.945***	7.109***	6.098***
		(2.679)	(5.460)	(3.883)	(4.513)
	Small Island	-9.783***	-11.409***	-19.557***	-18.049***
		(-2.507)	(-4.101)	(-4.143)	(-5.020)
	Religion Variables	Yes	Yes	Yes	Yes
	Residual k	0.025	0.010	-0.117*	-0.114**
		(0.490)	(0.235)	(-1.984)	(-2.382)
	Adj. R^2	0.60	0.59	0.35	0.38
	# Obs. (# Country)	64 (28)	75 (37)	82 (39)	93 (47)
$\frac{1-A}{A}$	Constant	3.531***	3.047	2.021*	2.138
		(4.195)	(1.202)	(1.857)	(1.047)
	α	-0.061***	-0.059***	-0.031***	-0.025**
		(-5.667)	(-3.384)	(-2.807)	(-2.056)
	k	0.005***	0.005***	0.004**	0.005**
		(2.808)	(2.778)	(2.075)	(2.199)
	Schooling 1960	0.260**	0.666***	0.445***	0.708***
		(2.265)	(7.608)	(4.012)	(7.903)
	Log Surface Area	-0.185**	-0.239	-0.129	-0.193
		(-2.211)	(-1.142)	(-1.429)	(-1.120)
	Landlocked	1.371**	1.329	-0.153	-0.237
		(1.947)	(1.109)	(-0.473)	(-0.659)
	Oil Prod.	2.916***	2.776***	2.737***	2.675***
		(5.786)	(5.972)	(5.033)	(5.724)
	Latitude	0.006	0.022	0.012	0.013
		(0.477)	(1.291)	(1.071)	(1.075)
	Residual k	-0.012*	-0.022***	-0.005	-0.009*
		(-1.650)	(-3.630)	(-0.981)	(-1.663)
	Adj. R^2	0.76	0.51	0.65	0.53
	# Obs. (# Country)	64 (28)	75 (37)	82 (39)	93 (47)
INEQ	Constant	60.517***	52.582***	56.693***	52.462***
		(12.638)	(8.650)	(12.627)	(11.017)
	α	-0.355***	-0.207*	-0.268**	-0.192*
		(-3.769)	(-1.888)	(-2.476)	(-1.878)
	k	-0.022	-0.009	-0.003	0.006
		(-1.306)	(-0.564)	(-0.190)	(0.362)
	$(1-A)/A$	-1.333***	-0.981**	-1.481***	-0.980**
		(-3.006)	(-2.340)	(-2.895)	(-2.166)
	Democracy	-5.271	-5.637**	-4.117	-6.345
		(-1.357)	(-1.254)	(-0.985)	(-1.562)
	Regional Dummies	Yes	Yes	Yes	Yes
	NET, PERSON, NA	Yes	Yes	Yes	Yes
	Residual k	0.071*	0.054	0.026	0.019
		(1.864)	(1.353)	(0.791)	(0.535)
	Adj. R^2	0.26	0.30	0.14	0.20
	# Obs. (# Country)	60 (27)	71 (37)	75 (39)	86 (47)

Notes: See the notes to above tables.

Table 5: Within-k for k, Base Sample

Dep. Var.	**RHS Variables**	**OLS** No Cont.	**OLS** with cont.	**GMM**
α	Constant	20.670***	19.853	19.852
		(13.133)	(0.843)	(0.909)
	Δk	-0.092*	-0.067	-0.067
		(-1.890)	(-1.355)	(-1.462)
	Arable land		0.065	0.065
			(1.067)	(1.150)
	Log Rural PopDen.		3.226**	3.226***
			(2.431)	(2.623)
	Small Island		-11.496***	-11.496***
			(-3.130)	(-3.377)
	Religion Variables	No	Yes	Yes
	Adj. R^2	0.02	0.56	0.56
	# Obs. (# Country)	64 (28)	64 (28)	64 (28)
$\frac{1-A}{A}$	Constant	3.726***	4.246***	4.5189**
		(6.726)	(4.876)	(5.625)
	α	-0.095***	-0.060***	-0.072***
		(-4.690)	(-5.225)	(-4.913)
	Δk	-0.008	-0.002	-0.003
		(-0.860)	(-0.301)	(-0.423)
	Schooling 1960		0.205*	0.163
			(1.662)	(1.456)
	Log Surface Area		-0.201**	-0.198**
			(-2.266)	(-2.373)
	Landlocked		0.988*	1.285**
			(1.736)	(2.041)
	Oil Prod.		2.786***	2.698***
			(4.974)	(5.148)
	Latitude		0.010	0.012
			(0.732)	(0.905)
	Adj. R^2	0.39	0.73	0.75
	# Obs. (# Country)	64 (28)	64 (28)	64 (28)
INEQ	Constant	56.345***	56.313***	54.231***
		(20.873)	(13.214)	(12.2184)
	α	-0.383***	-0.341***	-0.275**
		(-4.443)	(-3.430)	(-2.361)
	Δk	0.051*	0.037	0.042
		(1.647)	(1.138)	(1.370)
	$(1-A)/A$	-1.458***	-1.449***	-1.272***
		(-3.516)	(-3.171)	(-2.953)
	Democracy		-4.119	-4.001
			(-1.043))	(-1.083)
	Regional Dummies	No	Yes	Yes
	NET, PERSON, NA	Yes	Yes	Yes
	Adj. R^2	0.26	0.25	0.24
	# Obs. (# Country)	60 (27)	60 (27)	60 (27)

Notes: See the notes to above tables.

Table 6: ManVA/AgrVA for k, Base Sample, System Estimation (GMM)

Variables/Dep. Var.	α	$(1-A)/A$	$INEQ.$
Constant	67.156*** (5.394)	2.311*** (3.255)	55.381*** (12.923)
α		-0.017 (-1.137)	-0.249** (-2.102)
$\frac{ManVA}{AgrVA}$	-5.169*** (-6.534)	0.685*** (3.748)	0.513 (0.466)
$(1-A)/A$			-1.431** (-2.403)
Arable total land	0.074 (1.488)		
Log Rural Pop. Den.	0.588 (0.429)		
Small Island	-4.393** (-2.348)		
Schooling 1960		0.028 (0.403)	
Log Surface Area		-0.131* (-1.653)	
Landlocked		0.321 (0.825)	
Oil Producing		2.385*** (3.655)	
Latitude		0.001 (0.033)	
Democracy			-10.102** (-2.454)
Religion Var's.	Yes		
NET, PERSON, NA			Yes
Regional Dummies			Yes
Adj. R^2	0.70	0.81	0.38
# Obs. (# Country)	65 (29)	65 (29)	59 (29)

Notes: See the notes to above tables.

D Evolution of k

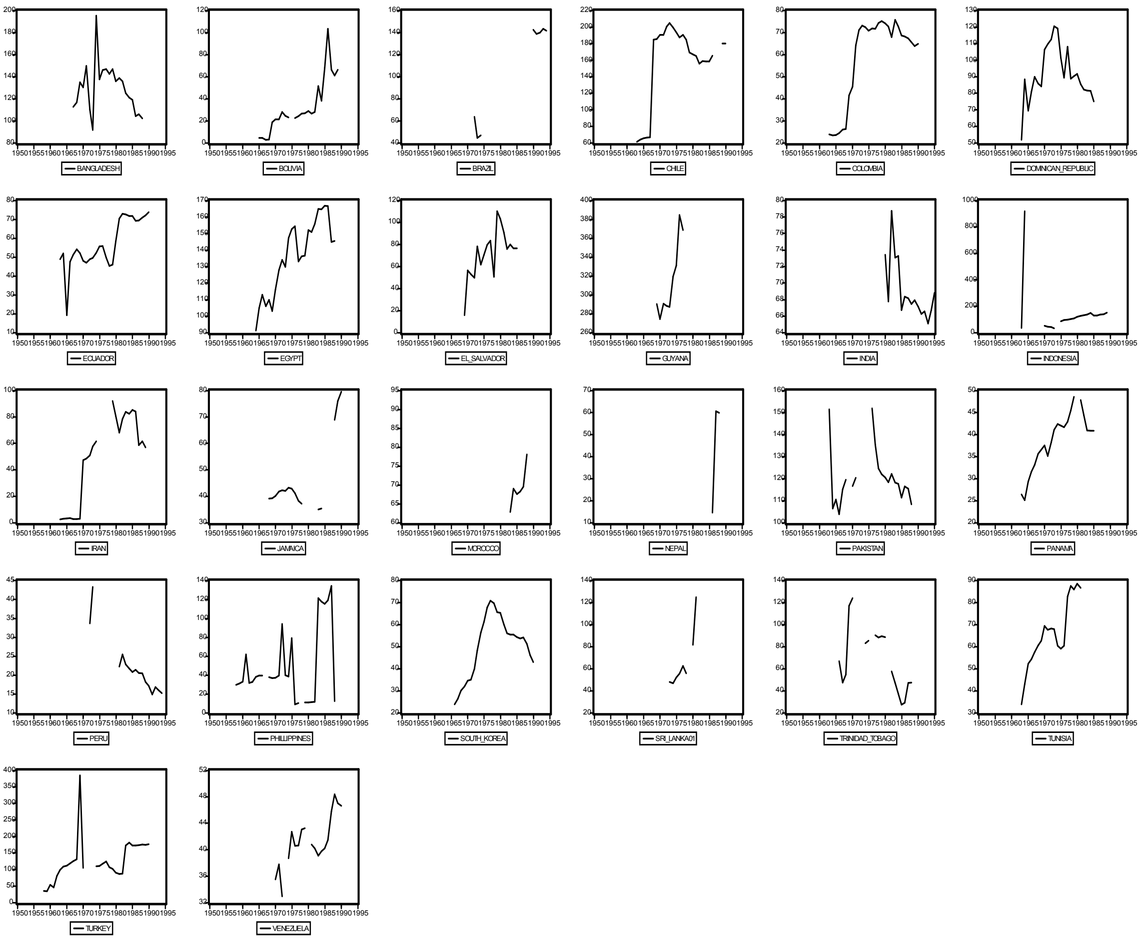

Figure 1. Developing Countries

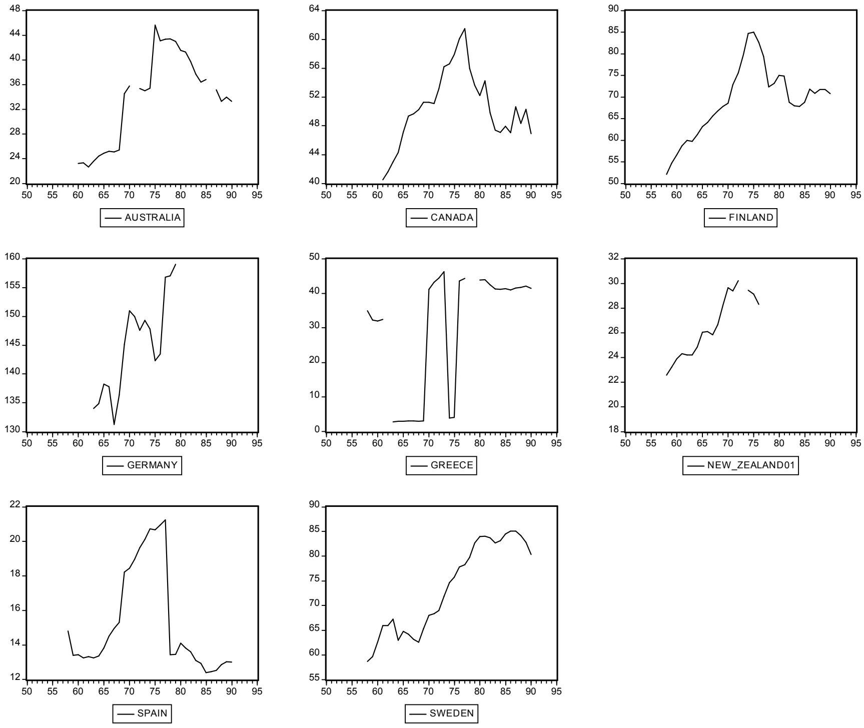

Figure 2. Developed Countries

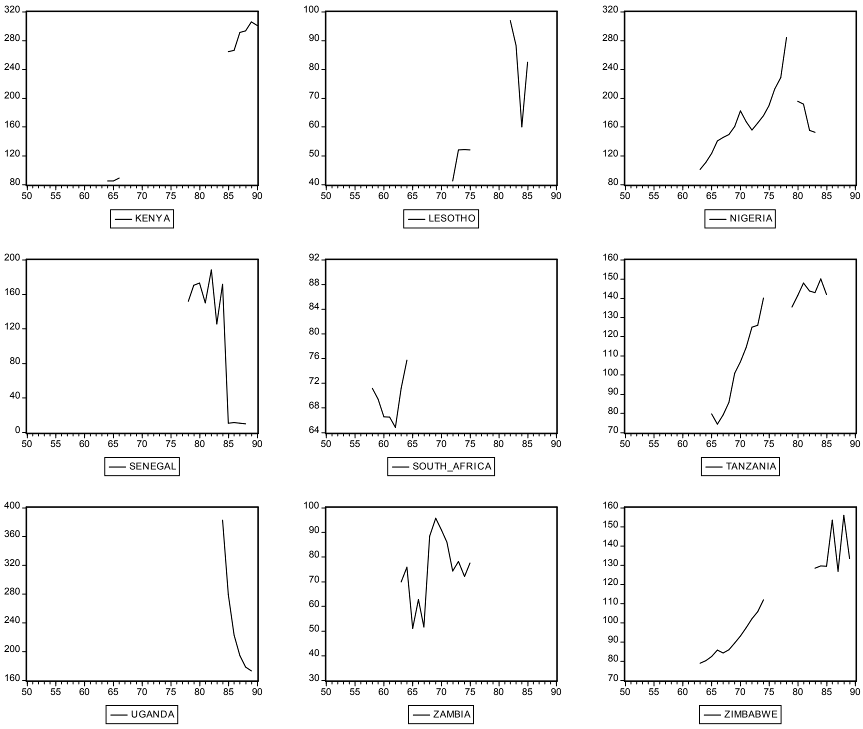

Figure 3. Sub-saharan African Countries